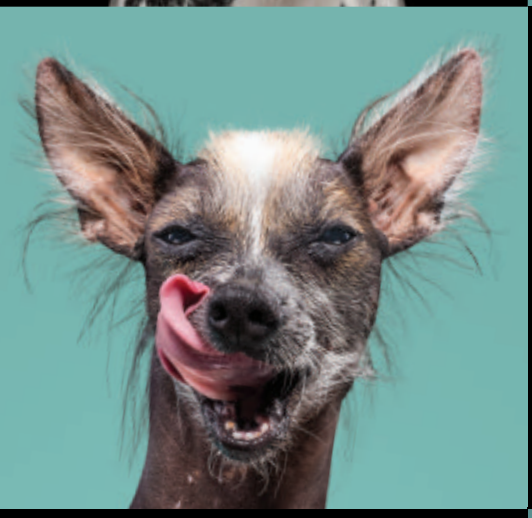
Avian influenza concerns for birds and humans pp. 441 & 459

Stalking dark matter particles with quantum sensors p. 448

Gamma-ray limits on nanohertz gravitational waves p. 521

Science

$15
29 APRIL 2022
science.org
AAAS



JUST LOOKS?

Breed predicts appearance more than behavior p. 475

2023 AWARDS & PRIZES

AAAS DAVID AND BETTY HAMBURG AWARD FOR SCIENCE DIPLOMACY

SUBMIT A NOMINATION TODAY
Deadline June 30, 2022

The **AAAS David and Betty Hamburg Award for Science Diplomacy** recognizes an individual or a limited number of individuals working together in the scientific and engineering or foreign affairs communities who are making an outstanding contribution to furthering science diplomacy.

Over the past 30 years, AAAS has honored an international cadre of science luminaries for their contributions to international scientific cooperation and science diplomacy. **In 2021, the award was renamed after David and Betty Hamburg** to recognize their unparalleled commitment to the significant role of science diplomacy to advance science, human rights, peace, and cooperation.

- The award is open to all regardless of nationality or citizenship.
- We accept self-nominations.
- Nominees must be living at the time of their nomination.

To learn more, visit

aaas.org/awards/science-diplomacy/about

AAAS gratefully acknowledges Carnegie Corporation of New York for their generous support to launch the AAAS David and Betty Hamburg Award for Science Diplomacy and the individuals and foundations whose contributions have begun an endowment that will allow us to sustain it in perpetuity.

PHOTO CREDIT: WWW.GOV.UK

2022 WINNER

Sir David King, Ph.D.

Former United Kingdom's Chief Scientific Advisor (2000-2007)

Honored for his diplomatic stewardship towards international consensus for urgent global action on climate change.

Winners will receive:

- Monetary prize of $10,000
- Commemorative plaque
- Worldwide promotion of their achievements through AAAS communication channels, including AAAS publications, member news, website, and social media
- Complimentary registration to the 2023 AAAS Annual Meeting
- Reimbursement for travel and hotel expenses to attend the AAAS Annual Meeting
- The opportunity to publish in *Science & Diplomacy*

AAAS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE

Your Legacy to Science

AN ESTATE GIFT TO THE
AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE

Since 1848, our founding year, the American Association for the Advancement of Science (AAAS) has been deeply committed to advancing science, engineering and innovation around the world for the benefit of all people.

By making AAAS a beneficiary of your will, trust, retirement plan or life insurance policy, you become a member of our 1848 Society, joining Thomas Edison, Alexander Graham Bell and the many distinguished individuals whose vision led to the creation of AAAS and our world-renowned journal, *Science*, so many years ago.

Unlike many of its peers, *Science* is not for-profit. Your estate gift would provide long-term financial stability and durable annual income that will support operations and competitive innovation for years to come. **This support is vital.**

"As a teacher and instructor, I bear responsibility for the younger generations. If you have extra resources, concentrate them on organizations, like AAAS, that are doing work for all."

—Prof. Elisabeth Ervin-Blankenheim, 1848 Society member

If you intend to include AAAS in your estate plans, provide this information to your lawyer or financial adviser:

Legal Name: American Association for the Advancement of Science
Federal Tax ID Number: 53-0196568
Address: 1200 New York Avenue, NW, Washington, DC 20005

If you would like more information on making an estate gift to AAAS, cut out and return the form below or send an email to philanthropy@aaas.org. Additional details are also available online at **www.aaas.org/1848Society**.

1848 SOCIETY
AAAS

cut here ✂

Yes, I would like more information about joining the AAAS 1848 Society.

PLEASE CONTACT ME AT:

Name: _____

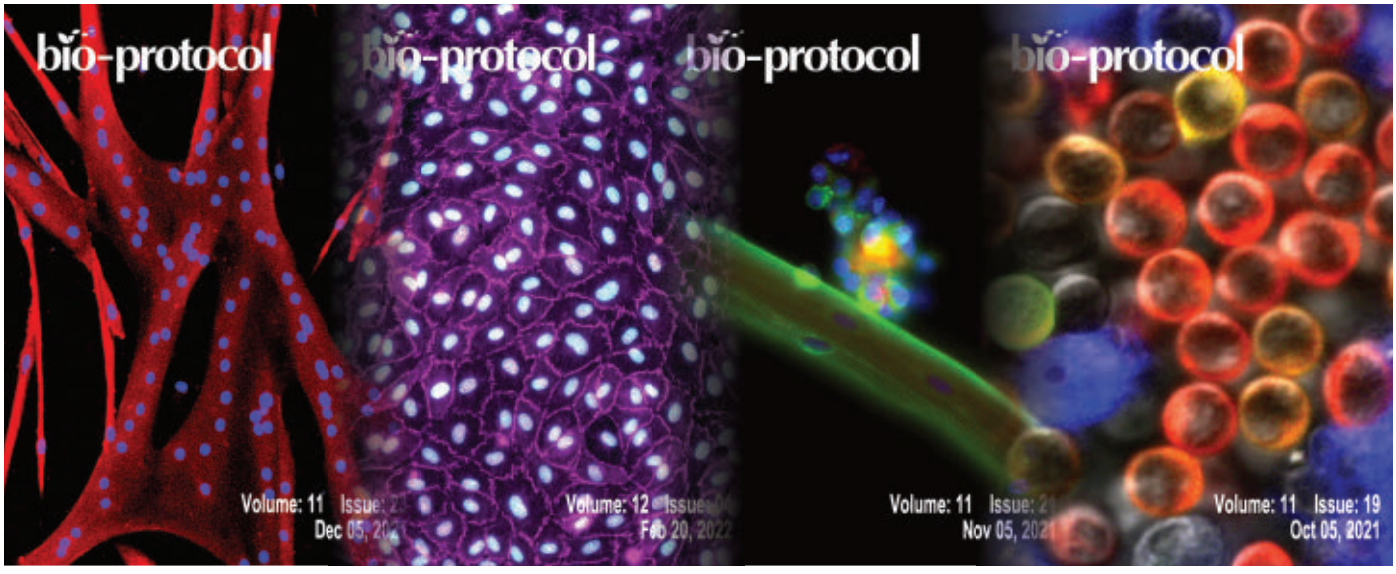
Address: _____

City: _____ State: _____ Zip code: _____ Country: _____

Email: _____ Phone: _____

RETURN THIS FORM TO:
AAAS Office of Philanthropy and Strategic Partnerships • 1200 New York Avenue, NW • Washington, DC 20005 USA

***Bio-protocol* is a strategic collaborator with *Science*/AAAS in support of scientific reproducibility.**



Bio-protocol journal covers

Taking on reproducibility with a team

As a grassroots effort, Bio-protocol *is building a community to foster interactions about the fine details of methodology in the life sciences.*

Reproducing results lies at the heart of life sciences, but it's a difficult challenge. "The complexity of what we do is so high, and there are so many variables that we try to control as much as we can, but that's the real reason why the reproducibility is so hard," says Ivan Zanoni, associate professor in pediatrics at Harvard Medical School and associate editor at *Bio-protocol*. As a journal—and now a major influencer in the scientific world—*Bio-protocol* is taking on that challenge.

"It just has to do with the facts of biology," says Holden Maecker, director of the Human Immune Monitoring Center at the Stanford University School of Medicine and advisor to *Bio-protocol*. "It's difficult to know which variables are going to be important for a given kind of experiment."

As a first step toward clarity and consistency, scientists need to communicate more about the precise details of a protocol. As then *Science* editor-in-chief Marcia McNutt wrote in 2014: "A transparent and rigorous approach … can almost always shine a light on issues of reproducibility" (*1*). The lack of such an approach spawned the *Bio-protocol* concept.

Addressing the problem

In 2010, as a postdoctoral researcher at Stanford University, bacterial expert Fanglian He started working on a new project, but she struggled to replicate methods from published research articles. She recalls, "It was very frustrating to spend time troubleshooting a method other people had already used successfully." In talks with colleagues, she learned that they were facing similar problems. So, in 2011, He launched *Bio-protocol*, where she is now the publisher.

This journal and its programs have a straightforward mission, as described by Caroline Shamu, associate dean for research cores and technology at Harvard Medical School and editor-in-chief of *Bio-protocol*: "It's a place to publish well-documented protocols relating to biology." Those protocols range from everyday methods to the latest trends. Plus, *Bio-protocol* continues the passion that led He to start this journal. Looking to help build a community rather than make a profit, *Bio-protocol* is free to read and includes more than 4,000 protocols.

Creating such a community depends on a broad-based interdisciplinary team, and associate editors make up a key component. In the summer and fall of 2021, the *Bio-protocol* editorial team convened a series of virtual check-in interviews with the associate editors. "It was great to see their enthusiasm to make improvements," Shamu says. "They made the effort to attend across many time zones."

To further expand *Bio-protocol*'s reach, it has formed strategic collaborations with other publications and organizations, including *eLife* and *Science*/AAAS.

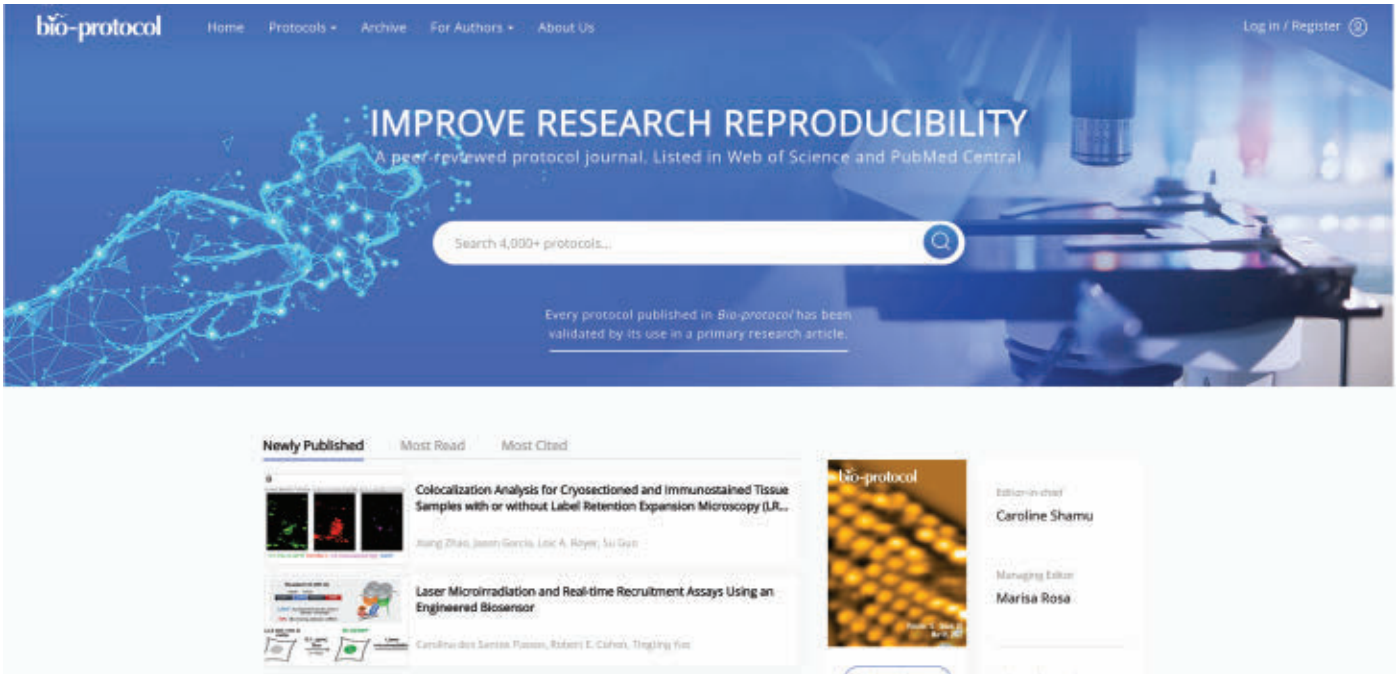
Putting the program to use

Any scientist who ever tried to replicate a protocol knows how difficult it can be without adequate information. "Open access to highly detailed protocols is essential for research," says Wolf B. Frommer, Alexander von Humboldt Professor at Heinrich-Heine-Universität in Düsseldorf and a member of the Leopoldina, the German National Academy of Sciences. "It saves an immense amount of effort and funds if protocols are as detailed as they are on the *Bio-protocol* website, since there are always little tricks or pitfalls that can be easily avoided with sufficient information."

A 2018 survey conducted by *Bio-protocol* shows that the plan is working. Of nearly 6,500 hundred scientists who downloaded protocols between October 2017 and January 2018, more than 1,500 responded to the survey. Of the approximately 700 scientists who had tried the protocols they had downloaded, 91% of them were able to reproduce the experiment. This means that scientists can turn to *Bio-protocol* to find tried-and-tested, reproducible methods that can be added to a lab's toolbox.

Bio-protocol also runs a sister database called Bio-101, which is an open platform for sharing protocols as well as asking and answering questions. "This is a very nice platform, which documents all of the everyday basic protocols—such

PHOTO: COURTESY OF BIO-PROTOCOL


bio-protocol
Home
Protocols
Archive
For Authors
About Us
Log in / Register
IMPROVE RESEARCH REPRODUCIBILITY
A peer-reviewed protocol journal. Listed in Web of Science and PubMed Central
Search 4,000+ protocols...
Every protocol published in Bio-protocol has been validated by its use in a primary research article.
Newly Published
Most Read
Most Cited
Colocalization Analysis for Cryosectioned and Immunostained Tissue Samples with or without Label Retention Expansion Microscopy (LR...
Laser Microirradiation and Real-time Recruitment Assays Using an Engineered Biosensor
Caroline Shamu
Marisa Rosa

The homepage of *Bio-protocol*

as polyacrylamide gel casting, or how to run a protein gel, or basic microscopy techniques," explains Amey Redkar, postdoctoral fellow at the University of Cordoba in Spain and associate editor at *Bio-protocol*. "If someone works under me as an intern, it's a good platform to give them the link to browse or search through for the kinds of techniques they might need in their projects." In addition to basic protocols, Bio-101 has expanded to include preprint protocols describing newly developed methods via the "Request a Protocol" initiative (*2*).

Sudhir Gopal Tattikota, a postdoctoral researcher in the lab of Norbert Perrimon at Harvard Medical School, sees clear benefits for researchers in Bio-protocol's Bio-101 online features. "Several readers simply approached us via the Bio-101 'Request a detailed protocol' option," he says, "and we were able to immediately send them detailed and stepwise protocols, which are sometimes difficult to incorporate within research articles."

In addition to helping readers, *Bio-protocol* benefits the authors as well. "It's great having a single, shared version of protocols," says Maecker. "It's made it easy for us to refer to those protocols in other papers and easy to share them with fellow scientists."

Plus, a *Bio-protocol* author must think very carefully and precisely about even the most basic elements of a method. "The idea is to write a protocol that is as detailed as possible—exactly which reagents, which type of plastic, which type of culture medium, which type of seal we used," says Zanoni. "It is very useful, because other people who read it can reproduce your data and your experiments, and you will be cited."

Growing the community

After more than a decade of working with life scientists around the world, the Bio-protocol organization plans to build an even bigger community of scientists dedicated to reproducibility. The basis of that work revolves around communication.

To make interactions even easier for life scientists, *Bio-protocol* includes a Q&A section at the end of each article. But communication can always be improved. "We continually strive for improvements that facilitate efficient communication between scientists," He says.

Improving that communication depends largely on the reach of *Bio-protocol*. One way that the organization hopes to expand its number of users is through its new Ambassador Program. As one of the ambassadors, Redkar describes it as a "program to promote this platform mainly to younger researchers—especially the undergrads who are in the lab for the first time." He adds, "It is just to make them aware of what *Bio-protocol* is and how it would benefit them in their everyday lab work."

The growing participation and engagement of authors and readers confirm the beneficial role that Bio-protocol—as an organization as well as a journal—is playing in the life sciences. As McNutt added in her editorial, exposing the issues of reproducibility "ensures that science moves forward, through independent verifications as well as the course corrections that come from refutations and the objective examination of the resulting data." That's just what He hoped for when she came up with the idea of *Bio-protocol*.

When scientists can reproduce a protocol, it helps everyone. As Zanoni says: "One of the best things is when another lab on the other side of the world can reproduce what you have done."

References

1. M. McNutt, *Science* **346**, 679 (2014); available at https://www.science.org/doi/10.1126/science.aaa1724.
2. "Request a Protocol (RaP)," Bio-101, *Bio-protocol*, https://bio-101.net/rap.

Sponsored by

bio-protocol
IMPROVE RESEARCH REPRODUCIBILITY

PHOTO: COURTESY OF BIO-PROTOCOL

CONTENTS

29 APRIL 2022 • VOLUME 376 • ISSUE 6592

448
A cloud of dark matter (predicted distribution in blue) surrounds colliding galaxy clusters.

NEWS

IN BRIEF

438 News at a glance

IN DEPTH

441 Deadly flu spreads through North American birds
As largest ever H5N1 outbreak hits poultry and wild species, researchers wonder whether virus is here to stay *By E. Stokstad*
PERSPECTIVE p. 459

442 India's speedy vaccine approvals come under fire
Critics say regulatory agency lacks key capabilities and independence *By P. Pulla*

444 NYU may hire biologist pushed out of MIT
Med school is in talks with David Sabatini, despite sexual misconduct findings *By M. Wadman*

445 Nicaragua's universities stagger under government pressure
Attacks on higher education autonomy put research—and researchers—at risk *By S. Moutinho*

446 First self-copying mRNA vaccine proves itself in pandemic trial
Twist on current vaccines reduces dose, eases distribution *By J. Cohen*

447 Math professor's trial next test of China Initiative
Mingqing Xiao is fighting U.S. charges that he failed to disclose ties to Chinese institutions *By J. Mervis*

FEATURES

448 A quantum sense for dark matter
By harnessing the strange rules of the subatomic realm, quantum sensors could solve one of the universe's biggest mysteries *By A. Cho*
PODCAST

INSIGHTS

PERSPECTIVES

452 A stark future for ocean life
Model predicts a mass extinction event in the oceans if climate change is uncurbed
By M. L. Pinsky and A. Fredston
REPORT p. 524

453 From outbreaks to endemic disease
Relief from the effects of epidemics may signal the start of low-level disease persistence *By M. F. Antolin*
REPORT p. 512

455 A granular approach to electrode design
The consistency of cathode particles plays a pivotal role in battery performance *By J. Xiao*
REPORT p. 517

466

456 Why do animals want what they like?
As in mammals, honey bee motivation for wanting rewards is modulated by dopamine
By J. E. Garcia and A. G. Dyer
REPORT p. 508

457 An adsorbent with flexible nanoscopic pores
Scanning transmission electron microscopy shows the adaptive pores of a zeolite
By T. Willhammar and X. Zou
RESEARCH ARTICLE p. 491

459 Resurgence of avian influenza virus
Unprecedented outbreaks of the H5N1 highly pathogenic avian influenza virus raise concern *By M. Wille and I. G. Barr*
NEWS STORY p. 441

461 Eugene N. Parker (1927–2022)
Pioneer of theoretical solar and plasma astrophysics *By S. D. Bale*

POLICY FORUM

462 A global system for the next generation of vaccines
COVID-19 has shown that hurdles can be overcome *By N. Arinaminpathy* et al.

BOOKS *ET AL.*

466 Landscapes of the Anthropocene
The late artist Per Kirkeby's preoccupation with geology is on display in a new exhibition
By D. Dixon

467 Rethinking the "Western" revolution in science
A historian sees global cultural and geopolitical roots in Europe's scientific breakthroughs *By J. Cañizares-Esguerra*

CREDITS: (TOP) (X-RAY) NASA/CXC/CFA/M. MARKEVITCH *ET AL.*; (OPTICAL) NASA/STSCI; MAGELLAN/U. ARIZONA/D. CLOWE *ET AL.*; (LENSING MAP) NASA/STSCI; ESO WFI; MAGELLAN/U. ARIZONA/D. CLOWE *ET AL.*; (BOTTOM) COURTESY OF MICHAEL WERNER GALLERY, NEW YORK AND LONDON

LETTERS

468 Retraction
By R. L. Reyes et al.

468 Evolutionary risks of osprey translocations
By F. Monti et al.

469 How to weaken Russian oil and gas strength
By R. Hausmann et al.

469 A global plastic treaty must cap production
By M. Bergmann et al.

470 Errata

RESEARCH

IN BRIEF

471 From *Science* and other journals

RESEARCH ARTICLES

474 Cancer
Stepwise-edited, human melanoma models reveal mutations' effect on tumor and microenvironment
E. Hodis et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABI8175

475 Dog genomics
Ancestry-inclusive dog genomics challenges popular breed stereotypes *K. Morrill* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABK0639

476 Cancer
Cancer cells use self-inflicted DNA breaks to evade growth limits imposed by genotoxic stress *B. D. Larsen* et al.

483 Neuroscience
Molecular and neural basis of pleasant touch sensation *B. Liu* et al.

491 Zeolites
In situ imaging of the sorption-induced subcell topological flexibility of a rigid zeolite framework *H. Xiong* et al.
PERSPECTIVE p. 457

REPORTS

496 Dynamic genome
Dynamics of CTCF- and cohesin-mediated chromatin looping revealed by live-cell imaging *M. Gabriele* et al.

502 Hydrogels
Piezoionic mechanoreceptors: Force-induced current generation in hydrogels
Y. Dobashi et al.

508 Neuroethology
Food wanting is mediated by transient activation of dopaminergic signaling in the honey bee brain *J. Huang* et al.
PERSPECTIVE p. 456

512 Epidemiology
Rabies shows how scale of transmission can enable acute infections to persist at low prevalence *R. Mancy* et al.
PERSPECTIVE p. 453; PODCAST

517 Batteries
Dynamics of particle network in composite battery cathodes *J. Li* et al.
PERSPECTIVE p. 455

521 Gravitational waves
A gamma-ray pulsar timing array constrains the nanohertz gravitational wave background
The Fermi-LAT Collaboration

524 Extinction
Avoiding ocean mass extinction from climate warming *J. L. Penn and C. Deutsch*
PERSPECTIVE p. 452

527 Organic chemistry
Scaffold hopping by net photochemical carbon deletion of azaarenes
J. Woo et al.

532 Organic chemistry
Accelerating reaction generality and mechanistic insight through additive mapping *C. N. Prieto Kullmer* et al.

DEPARTMENTS

437 Editorial
Removing a barrier to equity *By A. Mapp*

542 Working Life
Scientist and artist *By A. Bashir*

PHOTO: KIM TAYLOR/MINDEN PICTURES

ON THE COVER

Portraits of pet dogs exemplify the assortment of physical characteristics seen in purebred and mixed-breed dogs. Additionally, different breeds are often thought to have distinctive behaviors and personalities. However, data from thousands of dogs involved in the Darwin's Ark project reveal that this is rarely true, and that studying the full diversity of dogs is the best way to understand interactions between genetics, ancestry, and complex traits such as behavior. See page 475. *Photos: Elke Vogelsang*

Science Staff 436
Science Careers 540

SCIENCE (ISSN 0036-8075) is published weekly on Friday, except last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005. Periodicals mail postage (publication No. 484460) paid at Washington, DC, and additional mailing offices.

Copyright © 2022 by the American Association for the Advancement of Science. The title SCIENCE is a registered trademark of the AAAS.

Domestic individual membership, including subscription (12 months): $165 ($74 allocated to subscription). Domestic institutional subscription (51 issues): $2212; Foreign postage extra: Air assist delivery: $98. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request, GST #125488122. Publications Mail Agreement Number 1069624. **Printed in the U.S.A.**
Change of address: Allow 4 weeks, giving old and new addresses and 8-digit account number. **Postmaster:** Send change of address to AAAS, P.O. Box 96178, Washington, DC 20090–6178. **Single-copy sales:** $15 each plus shipping and handling available from backissues.science.org; bulk rate on request. **Authorization to reproduce** material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act can be obtained through the Copyright Clearance Center (CCC), www.copyright.com. The identification code for *Science* is 0036-8075. *Science* is indexed in the *Reader's Guide to Periodical Literature* and in several specialized indexes.

Science AAAS

Editor-in-Chief Holden Thorp, hthorp@aaas.org

Executive Editor Monica M. Bradford

Editors, Research Valda Vinson, Jake S. Yeston **Editor, Insights** Lisa D. Chong

DEPUTY EDITORS Stella M. Hurtley (UK), Phillip D. Szuromi, Sacha Vignieri **SR. EDITORIAL FELLOW** Andrew M. Sugden (UK) **SR. EDITORS** Gemma Alderton (UK), Caroline Ash (UK), Brent Grocholski, Pamela J. Hines, Di Jiang, Priscilla N. Kelly, Marc S. Lavine (Canada), Yevgeniya Nusinovich, Ian S. Osborne (UK), Beverly A. Purnell, L. Bryan Ray, H. Jesse Smith, Keith T. Smith (UK), Jelena Stajic, Peter Stern (UK), Valerie B. Thompson, Brad Wible, Yuen Yiu **ASSOCIATE EDITORS** Michael A. Funk, Bianca Lopez, Seth Thomas Scanlon (UK), Yury V. Suleymanov, Ekeoma Uzogara **LETTERS EDITOR** Jennifer Sills **LEAD CONTENT PRODUCTION EDITORS** Harry Jach, Lauren Kmec **CONTENT PRODUCTION EDITORS** Amelia Beyna, Jeffrey E. Cook, Chris Filiatreau, Julia Haber-Katris, Nida Masiulis, Abigail Shashikanth, Suzanne M. White **SR. EDITORIAL COORDINATORS** Carolyn Kyle, Beverly Shields **EDITORIAL COORDINATORS** Aneera Dobbins, Joi S. Granger, Jeffrey Hearn, Lisa Johnson, Maryrose Madrid, Ope Martins, Shannon McMahon, Jerry Richardson, Hilary Stewart (UK), Alice Whaley (UK), Anita Wynn **PUBLICATIONS ASSISTANTS** Alexander Kief, Ronmel Navas, Isabel Schnaidt, Brian White **EXECUTIVE ASSISTANT** Jessica Slater **ASI DIRECTOR, OPERATIONS** Janet Clements (UK) **ASI SR. OFFICE ADMINISTRATOR** Jessica Waldock (UK)

News Editor Tim Appenzeller

NEWS MANAGING EDITOR John Travis **INTERNATIONAL EDITOR** Martin Enserink **DEPUTY NEWS EDITORS** Elizabeth Culotta, Lila Guterman, David Grimm, Eric Hand (Europe), David Malakoff **SR. CORRESPONDENTS** Daniel Clery (UK), Jon Cohen, Jeffrey Mervis, Elizabeth Pennisi **ASSOCIATE EDITORS** Jeffrey Brainard, Kelly Servick, Catherine Matacic **NEWS REPORTERS** Adrian Cho, Jennifer Couzin-Frankel, Jocelyn Kaiser, Rodrigo Pérez Ortega (Mexico City), Robert F. Service, Erik Stokstad, Paul Voosen, Meredith Wadman **INTERN** Tess Joose **CONTRIBUTING CORRESPONDENTS** Warren Cornwall, Andrew Curry (Berlin), Ann Gibbons, Sam Kean, Eli Kintisch, Kai Kupferschmidt (Berlin), Andrew Lawler, Mitch Leslie, Eliot Marshall, Virginia Morell, Dennis Normile (Tokyo), Elisabeth Pain (Careers), Charles Piller, Gabriel Popkin, Michael Price, Joshua Sokol, Richard Stone, Emily Underwood, Gretchen Vogel (Berlin), Lizzie Wade (Mexico City) **CAREERS** Rachel Bernstein (Editor), Katie Langin (Associate Editor) **COPY EDITORS** Julia Cole (Senior Copy Editor), Morgan Everett, Cyra Master (Copy Chief) **ADMINISTRATIVE SUPPORT** Meagan Weiland

Creative Director Beth Rakouskas

DESIGN MANAGING EDITOR Marcy Atarod **GRAPHICS MANAGING EDITOR** Chris Bickel **INTERIM MULTIMEDIA MANAGING PRODUCER** Sarah Crespi **PHOTOGRAPHY MANAGING EDITOR** William Douthitt **WEB STRATEGY MANAGER** Kara Estelle-Powers **DESIGN EDITOR** Chrystal Smith **DESIGNER** Christina Aycock **GRAPHICS EDITOR** Nirja Desai **INTERACTIVE GRAPHICS EDITOR** Kelly Franklin **SENIOR GRAPHICS SPECIALISTS** Holly Bishop, Nathalie Cary **SENIOR SCIENTIFIC ILLUSTRATOR** Valerie Altounian **SCIENTIFIC ILLUSTRATOR** Ashley Mastin **VIDEO PRODUCER** Meagan Cantwell **SENIOR PHOTO EDITOR** Emily Petersen **PHOTO EDITOR** Kaitlyn Dolan **SOCIAL MEDIA STRATEGIST** Jessica Hubbard **SOCIAL MEDIA PRODUCER** Sabrina Jenkins **WEB DESIGNER** Jennie Pajerowski

Chief Executive Officer and Executive Publisher Sudip Parikh

Publisher, Science Family of Journals Bill Moran

DIRECTOR, BUSINESS SYSTEMS AND FINANCIAL ANALYSIS Randy Yi **DIRECTOR, BUSINESS OPERATIONS & ANALYSIS** Eric Knott **DIRECTOR OF ANALYTICS** Enrique Gonzales **MANAGER, BUSINESS OPERATIONS** Jessica Tierney **MANAGER, BUSINESS ANALYSIS** Cory Lipman **BUSINESS ANALYSTS** Kurt Ennis, Maggie Clark **FINANCIAL ANALYST** Isacco Fusi **BUSINESS OPERATIONS ADMINISTRATOR** Taylor Fisher **DIGITAL/PRINT STRATEGY MANAGER** Jason Hillman **SENIOR MANAGER, PUBLISHING AND CONTENT SYSTEMS** Marcus Spiegler **ASSISTANT MANAGER DIGITAL/PRINT** Rebecca Doshi **SENIOR CONTENT & PUBLISHING SYSTEMS SPECIALIST** Jacob Hedrick **SENIOR CONTENT SPECIALIST** Lori Murphy **PRODUCTION SPECIALIST** Kristin Wowk **DIGITAL PRODUCTION MANAGER** Lisa Stanford **CONTENT SPECIALIST** Kimberley Oster **ADVERTISING PRODUCTION OPERATIONS MANAGER** Deborah Tompkins **DESIGNER, CUSTOM PUBLISHING** Jeremy Huntsinger **SR. TRAFFIC ASSOCIATE** Christine Hall **SPECIAL PROJECTS ASSOCIATE** Sarah Dhere

ASSOCIATE DIRECTOR, BUSINESS DEVELOPMENT Justin Sawyers **GLOBAL MARKETING MANAGER** Allison Pritchard **DIGITAL MARKETING MANAGER** Aimee Aponte **JOURNALS MARKETING MANAGER** Shawana Arnold **MARKETING ASSOCIATES** Aaron Helmbrecht, Ashley Hylton, Mike Romano, Tori Velasquez, Jenna Voris **SENIOR DESIGNER** Kim Huynh

DIRECTOR AND SENIOR EDITOR, CUSTOM PUBLISHING Sean Sanders **ASSISTANT EDITOR, CUSTOM PUBLISHING** Jackie Oberst **PROJECT MANAGER** Melissa Collins

DIRECTOR, PRODUCT & PUBLISHING DEVELOPMENT Chris Reid **DIRECTOR, BUSINESS STRATEGY AND PORTFOLIO MANAGEMENT** Sarah Whalen **ASSOCIATE DIRECTOR, PRODUCT MANAGMENT** Kris Bishop **PRODUCT DEVELOPMENT MANAGER** Scott Chernoff **PUBLISHING TECHNOLOGY MANAGER** Michael Di Natale **SR. PRODUCT ASSOCIATE** Robert Koepke **PRODUCT ASSOCIATE** Caroline Breul, Anne Mason **SPJ ASSOCIATE MANAGER** Samantha Bruno Fuller **SPJ ASSOCIATE** Casey Buchta

MARKETING MANAGER Kess Knight **BUSINESS DEVELOPMENT MANAGER** Rasmus Andersen **SENIOR INSTITUTIONAL LICENSING MANAGER** Ryan Rexroth **INSTITUTIONAL LICENSING MANAGER** Marco Castellan, Claudia Paulsen-Young **SENIOR MANAGER, INSTITUTIONAL LICENSING OPERATIONS** Judy Lillibridge **SENIOR OPERATIONS ANALYST** Lana Guz **SYSTEMS & OPERATIONS ANALYST** Ben Teincuff **FULFILLMENT ANALYST** Aminta Reyes

DIRECTOR, GLOBAL SALES Tracy Holmes **US EAST COAST AND MID WEST SALES** Stephanie O'Connor **US MID WEST, MID ATLANTIC AND SOUTH EAST SALES** Chris Hoag **US WEST COAST SALES** Lynne Stickrod **ASSOCIATE DIRECTOR, ROW** Roger Goncalves **SALES REP, ROW** Sarah Lelarge **SALES ADMIN ASSISTANT, ROW** Victoria Glasbey **DIRECTOR OF GLOBAL COLLABORATION AND ACADEMIC PUBLISHING RELATIONS, ASIA** Xiaoying Chu **ASSOCIATE DIRECTOR, INTERNATIONAL COLLABORATION** Grace Yao **SALES MANAGER** Danny Zhao **MARKETING MANAGER** Kilo Lan **ASCA CORPORATION, JAPAN** Yoshimi Toda (Tokyo), Miyuki Tani (Osaka)

DIRECTOR, COPYRIGHT, LICENSING AND SPECIAL PROJECTS Emilie David **RIGHTS AND PERMISSIONS ASSOCIATE** Elizabeth Sandler **LICENSING ASSOCIATE** Virginia Warren **CONTRACT SUPPORT SPECIALIST** Michael Wheeler

MAIN HEADQUARTERS
Science/AAAS
1200 New York Ave. NW
Washington, DC 20005

SCIENCE INTERNATIONAL
Clarendon House
Clarendon Road
Cambridge, CB2 8FH, UK

SCIENCE CHINA
Room 1004, Culture Square
No. 59 Zhongguancun St.
Haidian District, Beijing, 100872

SCIENCE JAPAN
ASCA Corporation
Sibaura TY Bldg. 4F, 1-14-5
Shibaura Minato-ku
Tokyo, 108-0073 Japan

EDITORIAL
science_editors@aaas.org

NEWS
science_news@aaas.org

INFORMATION FOR AUTHORS
science.org/authors/
science-information-authors

REPRINTS AND PERMISSIONS
science.org/help/
reprints-and-permissions

MEDIA CONTACTS
scipak@aaas.org

MULTIMEDIA CONTACTS
SciencePodcast@aaas.org
ScienceVideo@aaas.org

INSTITUTIONAL SALES AND SITE LICENSES
science.org/librarian

PRODUCT ADVERTISING & CUSTOM PUBLISHING
advertising.science.org/
products-services
science_advertising@aaas.org

CLASSIFIED ADVERTISING
advertising.science.org/
science-careers
advertise@sciencecareers.org

JOB POSTING CUSTOMER SERVICE
employers.sciencecareers.org
support@sciencecareers.org

MEMBERSHIP AND INDIVIDUAL SUBSCRIPTIONS
science.org/subscriptions

MEMBER BENEFITS
aaas.org/membership/benefits

AAAS BOARD OF DIRECTORS
CHAIR Claire M. Fraser
PRESIDENT Susan G. Amara
PRESIDENT-ELECT Gilda A. Barabino
TREASURER Carolyn N. Ainslie
CHIEF EXECUTIVE OFFICER Sudip Parikh
BOARD Cynthia M. Beall, Rosina M. Bierbaum, Ann Bostrom, Janine Austin Clayton, Laura H. Greene, Kaye Husbands Fealing, Maria M. Klawe, Robert B. Millard, William D. Provine

BOARD OF REVIEWING EDITORS (Statistics board members indicated with **(S)**)

Erin Adams, *U. of Chicago*
Takuzo Aida, *U. of Tokyo*
Leslie Aiello, *Wenner-Gren Fdn.*
Deji Akinwande, *UT Austin*
Judith Allen, *U. of Manchester*
Marcella Alsan, *Harvard U.*
Sebastian Amigorena, *Inst. Curie*
James Analytis, *UC Berkeley*
Trevor Archer, *NIEHS, NIH*
Paola Arlotta, *Harvard U.*
David Awschalom, *U. of Chicago*
Delia Baldassarri, *NYU*
Nenad Ban, *ETH Zürich*
Nandita Basu, *U. of Waterloo*
Franz Bauer, *Pontificia U. Católica de Chile*
Ray H. Baughman, *UT Dallas*
Carlo Beenakker, *Leiden U.*
Yasmine Belkaid, *NIAID, NIH*
Philip Benfey, *Duke U.*
Kiros T. Berhane, *Columbia U.*
Joseph J. Berry, *NREL*
Alessandra Biffi, *Harvard Med.*
Chris Bowler, *École Normale Supérieure*
Ian Boyd, *U. of St. Andrews*
Malcolm Brenner, *Baylor Coll. of Med.*
Emily Brodsky, *UC Santa Cruz*
Ron Brookmeyer, *UCLA* **(S)**
Christian Büchel, *UKE Hamburg*
Dennis Burton, *Scripps Res.*
Carter Tribley Butts, *UC Irvine*
György Buzsáki, *NYU School of Med.*
Mariana Byndloss, *Vanderbilt U. Med. Ctr.*
Annmarie Carlton, *UC Irvine*
Simon Cauchemez, *Inst. Pasteur*
Ling-Ling Chen, *SIBCB, CAS*
M. Keith Chen, *UCLA*
Zhijian Chen, *UT Southwestern Med. Ctr.*
Ib Chorkendorff, *Denmark TU*
Amander Clark, *UCLA*
James J. Collins, *MIT*
Robert Cook-Deegan, *Arizona State U.*
Virginia Cornish *Columbia U.*
Carolyn Coyne, *Duke U.*
Roberta Croce, *VU Amsterdam*
Christina Curtis, *Stanford U.*
Ismaila Dabo, *Penn State U.*
Jeff L. Dangl, *UNC*
Chiara Daraio, *Caltech*
Nicolas Dauphas, *U. of Chicago*
Christian Davenport, *U. of Michigan*
Frans de Waal, *Emory U.*
Claude Desplan, *NYU*
Sandra Díaz, *U. Nacional de Córdoba*
Samuel Díaz-Muñoz, *UC Davis*
Ulrike Diebold, *TU Wien*
Stefanie Dimmeler, *Goethe-U. Frankfurt*
Hong Ding, *Inst. of Physics, CAS*
Dennis Discher, *UPenn*
Jennifer A. Doudna, *UC Berkeley*
Ruth Drdla-Schutting, *Med. U. Vienna*
Raissa M. D'Souza, *UC Davis*
Bruce Dunn, *UCLA*
William Dunphy, *Caltech*
Scott Edwards, *Harvard U.*
Todd Ehlers, *U. of Tübingen*
Nader Engheta, *UPenn*
Karen Ersche, *U. of Cambridge*
Beate Escher, *UFZ & U. of Tübingen*
Barry Everitt, *U. of Cambridge*
Vanessa Ezenwa, *U. of Georgia*
Toren Finkel, *U. of Pitt. Med. Ctr.*
Gwenn Flowers, *Simon Fraser U.*
Peter Fratzl, *Max Planck Inst. Potsdam*
Elaine Fuchs, *Rockefeller U.*
Jay Gallagher, *U. of Wisconsin*
Daniel Geschwind, *UCLA*
Ramon Gonzalez, *U. of South Florida*
Sandra González-Bailón, *UPenn*
Nicolas Gruber, *ETH Zürich*
Hua Guo, *U. of New Mexico*
Taekjip Ha, *Johns Hopkins U.*
Daniel Haber, *Mass. General Hos.*
Sharon Hammes-Schiffer, *Yale U.*
Wolf-Dietrich Hardt, *ETH Zürich*
Louise Harra, *U. Coll. London*
Carl-Philipp Heisenberg, *IST Austria*
Janet G. Hering, *Eawag*
Christoph Hess, *U. of Basel & U. of Cambridge*
Heather Hickman, *NIAID, NIH*
Hans Hilgenkamp, *U. of Twente*
Janneke Hille Ris Lambers, *ETH Zürich*
Kai-Uwe Hinrichs, *U. of Bremen*
Deirdre Hollingsworth, *U. of Oxford*
Randall Hulet, *Rice U.*
Auke Ijspeert, *EPFL*
Gwyneth Ingram, *ENS Lyon*
Darrell Irvine, *MIT*
Akiko Iwasaki, *Yale U.*
Stephen Jackson, *USGS & U. of Arizona*
Erich Jarvis, *Rockefeller U.*
Peter Jonas, *IST Austria*
Matt Kaeberlein, *U. of Wash.*
William Kaelin Jr., *Dana-Farber Cancer Inst.*
Daniel Kammen, *UC Berkeley*
Kisuk Kang, *Seoul Nat. U.*
Sabine Kastner, *Princeton U.*
V. Narry Kim, *Seoul Nat. U.*
Robert Kingston, *Harvard Med.*
Nancy Knowlton, *Smithsonian Institution*
Etienne Koechlin, *École Normale Supérieure*
Alex L. Kolodkin, *Johns Hopkins U.*
Julija Krupic, *U. of Cambridge*
Paul Kubes, *U. of Calgary*
Gabriel Lander, *Scripps Res.* **(S)**
Mitchell A. Lazar, *UPenn*
Luis Liz-Marzán, *CIC biomaGUNE*
Omar Lizardo, *UCLA*
Jonathan Losos, *Wash. U. in St. Louis*
Ke Lu, *Inst. of Metal Res., CAS*
Christian Lüscher, *U. of Geneva*
Jean Lynch-Stieglitz, *Georgia Inst. of Tech.*
David Lyons, *U. of Edinburgh*
Fabienne Mackay, *QIMR Berghofer*
Anne Magurran, *U. of St. Andrews*
Ari Pekka Mähönen, *U. of Helsinki*
Asifa Majid, *U. of York*
Oscar Marín, *King's Coll. London*
Charles Marshall, *UC Berkeley*
Christopher Marx, *U. of Idaho*
David Masopust, *U. of Minnesota*
Geraldine Masson, *CNRS*
C. Robertson McClung, *Dartmouth*
Rodrigo Medellín, *U. Nacional Autónoma de México*
C. Jessica Metcalf, *Princeton U.*
Baoxia Mi, *UC Berkeley*
Tom Misteli, *NCI, NIH*
Alison Motsinger-Reif, *NIEHS, NIH* **(S)**
Suresh Naidu, *Columbia U.*
Danielle Navarro, *U. of New South Wales*
Daniel Nettle, *Newcastle U.*
Daniel Neumark, *UC Berkeley*
Beatriz Noheda, *U. of Groningen*
Helga Nowotny, *Vienna Sci. & Tech. Fund*
Rachel O'Reilly, *U. of Birmingham*
Pilar Ossorio, *U. of Wisconsin*
Andrew Oswald, *U. of Warwick*
Isabella Pagano, *Istituto Nazionale di Astrofisica*
Elizabeth Levy Paluck, *Princeton U.*
Jane Parker, *Max Planck Inst. Cologne*
Giovanni Parmigiani, *Dana-Farber Cancer Inst.* **(S)**
Daniel Pauly, *U. of British Columbia*
Ana Pêgo, *U. do Porto*
Samuel Pfaff, *Salk Inst.*
Julie Pfeiffer, *UT Southwestern Med. Ctr.*
Philip Phillips, *UIUC*
Matthieu Piel, *Inst. Curie*
Kathrin Plath, *UCLA*
Martin Plenio, *Ulm U.*
Katherine Pollard, *UCSF*
Elvira Poloczanska, *Alfred-Wegener-Inst.*
Julia Pongratz, *Ludwig Maximilians U.*
Philippe Poulin, *CNRS*
Jonathan Pritchard, *Stanford U.*
Lei Stanley Qi, *Stanford U.*
Trevor Robbins, *U. of Cambridge*
Joeri Rogelj, *Imperial Coll. London*
Amy Rosenzweig, *Northwestern U.*
Mike Ryan, *UT Austin*
Miquel Salmeron, *Lawrence Berkeley Nat. Lab*
Nitin Samarth, *Penn State U.*
Erica Ollmann Saphire, *La Jolla Inst.*
Joachim Saur, *U. zu Köln*
Alexander Schier, *Harvard U.*
Wolfram Schlenker, *Columbia U.*
Susannah Scott, *UC Santa Barbara*
Anuj Shah, *U. of Chicago*
Vladimir Shalaev, *Purdue U.*
Jie Shan, *Cornell U.*
Beth Shapiro, *UC Santa Cruz*
Jay Shendure, *U. of Wash.*
Steve Sherwood, *U. of New South Wales*
Brian Shoichet, *UCSF*
Robert Siliciano, *JHU School of Med.*
Lucia Sivilotti, *U. Coll. London*
Richard Smith, *UNC* **(S)**
John Speakman, *U. of Aberdeen*
Tara Spires-Jones, *U. of Edinburgh*
Allan C. Spradling, *Carnegie Institution for Sci.*
V. S. Subrahmanian, *Northwestern U.*
Ira Tabas, *Columbia U.*
Eriko Takano, *U. of Manchester*
Patrick Tan, *Duke-NUS Med. School*
Sarah Teichmann, *Wellcome Sanger Inst.*
Rocio Titiunik, *Princeton U.*
Shubha Tole, *Tata Inst. of Fundamental Res.*
Maria-Elena Torres Padilla, *Helmholtz Zentrum München*
Kimani Toussaint, *Brown U.*
Barbara Treutlein, *ETH Zürich*
Jason Tylianakis, *U. of Canterbury*
Wim van der Putten, *Netherlands Inst. of Ecology*
Ivo Vankelecom, *KU Leuven*
Judith Varner, *UC San Diego*
Henrique Veiga-Fernandes, *Champalimaud Fdn.*
Reinhilde Veugelers, *KU Leuven*
Bert Vogelstein, *Johns Hopkins U.*
Julia Von Blume, *Yale School of Med.*
David Wallach, *Weizmann Inst.*
Jane-Ling Wang, *UC Davis* **(S)**
Jessica Ware, *Amer. Mus. of Natural Hist.*
David Waxman, *Fudan U.*
Chris Wikle, *U. of Missouri* **(S)**
Terrie Williams, *UC Santa Cruz*
Ian A. Wilson, *Scripps Res.* **(S)**
Hao Wu, *Harvard U.*
Li Wu, *Tsinghua U.*
Wei Xie, *Tsinghua U.*
Benjamin Youngblood, *St. Jude*
Yu Xie, *Princeton U.*
Jan Zaanen, *Leiden U.*
Kenneth Zaret, *UPenn School of Med.*
Bing Zhu, *Inst. of Biophysics, CAS*
Xiaowei Zhuang, *Harvard U.*
Maria Zuber, *MIT*

Science serves as a forum for discussion of important issues related to the advancement of science by publishing material on which a consensus has been reached as well as including the presentation of minority or conflicting points of view. Accordingly, all articles published in *Science*—including editorials, news and comment, and book reviews—are signed and reflect the individual views of the authors and not official points of view adopted by AAAS or the institutions with which the authors are affiliated.

EDITORIAL

Removing a barrier to equity

Earlier this year, the University of Michigan became the first US university to remove the requirement that applicants to its nonprofessional doctoral programs take a standardized test—the Graduate Record Examination (GRE). This decision will not, on its own, address inequities in admissions practice, nor the broader education barriers that many applicants face. But it is a major step toward an admissions process that considers all dimensions of a candidate's preparation and promise—a holistic view that should be adopted by all universities if equity in education and opportunities is to be achieved.

The stated goal of the GRE general test is to assess writing skills and quantitative and verbal reasoning. Its actual value as an instrument for determining who is likely to be successful in a doctoral graduate program, however, has not been established. There is a positive correlation across disciplines between GRE scores and first-year graduate course GPA (grade point average). However, although coursework is part of obtaining foundational knowledge in a field, the central focus of a doctoral degree comprises intellectual contributions and research accomplishments (measured in many dimensions, such as publications, presentations, and patents).

> "...a truly holistic admissions process... requires more than dropping standardized tests."

Beyond major questions about GRE utility, there are substantive financial and opportunity costs associated with requiring the GRE: costs for applicants and costs for admissions committees. For the potential applicant, the test is expensive; registration alone has a price tag of US$205, with additional fees for score reporting. (A fee reduction is possible for qualified applicants.) Many feel pressure to enroll in expensive GRE prep courses that teach test-taking strategies and often offer money-back guarantees for competitive scores and score improvements. The GRE 162+ course offered by Princeton Review, for example, guarantees scores of at least 162 (out of 170) at a cost of US$2149.

What are the costs for admissions committees that use the GRE in admissions decisions? In short, the loss of talented applicants at every stage of the process. Students who do not have the financial means to prepare for and/or take the test will not apply. A further loss comes from potential applicants who lack access to the test, either because of the limited availability of physical testing sites or because of lack of access to wireless networks needed for online test-taking options. Once applications have arrived for review by a doctoral admissions committee, use of the GRE can lead to additional loss of talented applicants. Despite recommendations that GRE scores only be used in the context of an overall evaluation of an applicant, in practice they can readily be employed—implicitly or explicitly—as cutoffs for further consideration of an application. This is an especially problematic practice given that GRE scores consistently correlate with the sex, ethnicity, socioeconomic status, and race of the test taker.

Discontinuing the use of the GRE in doctoral admissions is, therefore, a valuable step toward best practices to reduce the barriers—systemic and otherwise—faced by students, to decrease bias in the evaluation process, and to support admission and recruitment of outstanding student cohorts. A concern is that in the absence of the GRE score, admissions committees will rely more heavily on other aspects of a candidate's application that—like the GRE—are strongly correlated with race, sex, and socioeconomic factors rather than talent and promise. This includes undergraduate institution ranking, prestigious internships, and performance on other standardized tests. It is for exactly this reason that frank conversations about discordance between the stated goals and values of doctoral programs and the common practices in doctoral admissions are indispensable.

In recent years, many individual graduate programs have removed the GRE requirement, with notable success. But implementing a truly holistic admissions process at the university level requires more than dropping standardized tests. Faculty need education and institutional support to develop effective and equitable evaluation rubrics and to learn how to better use the rich information in graduate applications rather than single data points in their evaluative processes. The continual collection and analysis of data surrounding all aspects of the student life cycle—from recruitment through graduation—are also critical. In this way, changes across many dimensions that correlate with GRE general test discontinuation and evolving holistic admissions practices can be identified and can inform future goals for the graduate education community at large.

–Anna Mapp

Anna Mapp
is the Edwin Vedejs Collegiate Professor of Chemistry in the Department of Chemistry, Research Professor in the Life Sciences Institute, and Associate Dean of Biological and Health Sciences at the Rackham Graduate School, University of Michigan, Ann Arbor, MI, USA.
amapp@umich.edu

PHOTO: STEPHANIE KING

10.1126/science.abq6714

NEWS

> “The reality is that there is a loss of momentum.”
>
> **Isaac Adewole,** a physician and consultant for the Africa Centres for Disease Control and Prevention, in *The New York Times*, on data showing the world is falling far short of a goal of vaccinating 70% of adults against COVID-19. Developing countries are more likely to have low rates.

IN BRIEF

Edited by **Jeffrey Brainard**

A cliff marks the edge of an ancient delta, where the Perseverance rover will collect rocks for testing.

PLANETARY SCIENCE

Rover reaches Mars river delta

NASA's Perseverance rover last week reached the primary target of its mission, an ancient river delta, a fan-shaped collection of rocks and sediments, frozen in time billions of years ago, that spills from the rim of Jezero crater to its floor. Operators plan to comb the fossilized delta sediments for rocks that could contain organic molecules indicative of past martian life. The $2.7 billion rover landed on Mars in February 2021, studying volcanic rocks on the crater floor for a year before trundling to the delta, which ends in a cliff face at the crater's rim. The rover's team plans to drill and collect approximately 30 rock samples, which are expected to be returned to Earth in future missions that NASA and the European Space Agency plan to launch in 2028.

Atom smasher powers up again

PARTICLE PHYSICS | The world's largest atom smasher, the Large Hadron Collider (LHC) at CERN, the European particle physics laboratory, started up last week after more than 3 years of upgrades. The 27-kilometer-long accelerator famously blasted out the long-sought Higgs boson in 2012, and it last took data in December 2018. Since then, technicians have improved the accelerators that feed protons into countercirculating beams and the four large detectors that study the resulting proton collisions. For the next few months, researchers will work out the bugs and slowly increase the beams' energy. Data recording—and the search for new and unexpected particles—should begin this summer. In 2025, the LHC will shut down for an even more extensive upgrade to greatly increase the intensity of its beams, which would increase the rate of proton collisions.

Biden resets environmental rules

POLICY | Reversing decisions made by former President Donald Trump, the administration of President Joe Biden last week released new guidelines expanding the technical issues that must be assessed in environmental reviews of a wide range of federally funded activities, including major construction projects. Trump had narrowed the scope of the reviews required by the National Environmental Policy Act (NEPA). For example, federal agencies were directed to consider only the direct, near-term effects of an activity, meaning they could ignore longer term issues, such as a project's impact on future climate change. The Biden administration says it plans to release a second set of NEPA guideline revisions that will address other controversial alterations made by Trump, including tight time limits on reviews.

Restore degraded land, U.N. urges

SUSTAINABILITY | Reversing global land degradation can help alleviate three big problems—the effects of climate change, biodiversity loss, and food insecurity, a U.N. report says. The *Global Land Outlook 2*, released on 27 April by the United Nations Convention to Combat Desertification, quantifies the toll taken by development, deforestation, agriculture, and other human activities on soil health. If current trends continue, the area of degraded land will equal the size of South America and global crop and ecosystem productivity will drop as much as 14% by 2050. Sub-Saharan Africa will fare the worst. But through steps such as planting trees and stabilizing erosion on 5 billion hectares, countries can slow this decline, increase crop yields, and lock up more atmospheric carbon in soils. Already more than 100 nations have pledged to restore 1 billion hectares by 2030. Redirecting money spent on fossil fuel and farming subsidies can help pay for restoration, the report says. It says restoration means not just creating protected wildlands, but also taking active steps to enable land to support farming, tree plantations, and grazing.

PHOTO: NASA/JPL-CALTECH/ASU

PARTICLE PHYSICS

Fermilab neutrino source gets green light for construction

Fermi National Accelerator Laboratory (Fermilab) last week won final approval to build a new superconducting linear accelerator that will generate an intense beam of protons with an energy of 800 million electron volts. Researchers will use that beam to create elusive particles called neutrinos and shoot them 1300 kilometers through Earth at a planned underground detector in South Dakota for high-priority experiments.

The new accelerator, the Proton Improvement Plan-II, replaces an older one with half the energy. Workers had already begun to construct buildings for the new machine; last week's go-ahead from the U.S. Department of Energy lets the lab build the actual accelerator. All told, the project will cost $978 million, not including equipment contributions from other countries totaling roughly $330 million. Completion is expected in 2032.

A new U.S. accelerator will speed up particles using a series of these five-cell, superconducting, radio-frequency cavities.

Vikings sold ivory in Ukraine

ARCHAEOLOGY | Vikings shipped walrus ivory from Greenland far away to Kyiv, according to new analyses of ancient walrus skulls and ivory figurines discovered in the Ukrainian capital. The 4000-kilometer trade route is much longer than was thought based on previous studies, which suggested eastern users of the coveted material sourced it only from walruses caught in the Russian arctic. After discovering finished ivory artifacts and discarded walrus skull fragments in a Kyiv layer dating to the 1100s C.E., an international team used DNA and chemical analyses to link the material to a genetic group of walruses found only in the western Atlantic Ocean. Vikings appear to have overhunted the Greenlandic walruses, and the authors say consumer demand in Eastern Europe may have been responsible. A decline in harvest could also help explain why Norse settlers abandoned Greenland in the 1300s.

Walrus ivory, used in this carved specimen from Norway, was widely prized in early medieval Europe.

PHOTOS: (TOP TO BOTTOM) REIDAR HAHN/FERMILAB; ÅGE HOJEM/NTNU UNIVERSITY MUSEUM

Asteroid explorer gets 2nd target

PLANETARY SCIENCE | Next year, after swinging past Earth and dropping off a cargo of rocks it collected from the asteroid Bennu, NASA's OSIRIS-REx spacecraft will embark on a new mission to the asteroid Apophis, the agency said this week. In 2029, the 300-meter-wide Apophis will zip past Earth in the closest flyby of a large asteroid in modern records; at one-tenth the distance to the Moon, it will be visible to the naked eye. Afterward, the renamed OSIRIS-Apex mission will visit and collect data from the asteroid for 18 months. Although the spacecraft cannot collect any more rocks, it will gauge how Apophis's orbit was perturbed by Earth's gravity and conduct an approach maneuver to scorch off the asteroid's surface layer, exposing hidden layers below.

Most in U.S. show signs of infection

COVID-19 | Nearly three in five U.S. residents had been infected with SARS-CoV-2 by February, a sign of the speed with which the highly contagious Omicron variant surged, the Centers for Disease Control and Prevention reported this week. Starting in September 2021, researchers analyzed tens of thousands of blood samples collected each month, measuring antibodies that are generated by exposure to the virus but not by the vaccines available in the United States. From December 2021 to February, as Omicron took off, the portion of samples with the antibodies rose from 34% to 58% across all age groups. In children and adolescents, the least vaccinated age groups, the proportion rose from 45% to 75%, the researchers reported on 26 April in the *Morbidity and Mortality Weekly Report*. Other studies have shown that COVID-19 vaccines protect most recipients against severe illness, but effectiveness against infection, especially by Omicron, wanes.

Avoidable COVID-19 deaths tallied

PUBLIC HEALTH | About 234,000 U.S. deaths from COVID-19 since June 2021 could have been prevented if all eligible adults had received the primary series of vaccinations, an analysis by the Kaiser Family Foundation

estimated. The toll represents 60% of all adult COVID-19 deaths since then and one-quarter of the nearly 1 million U.S. deaths from COVID-19 since 2020.

Pandemic cut kids' vax rates

INFECTIOUS DISEASES | COVID-19 drove down the rate at which U.S. kindergartners received three routine vaccinations during the 2020–21 school year, the Centers for Disease Control and Prevention said last week. In 47 states, percentages of those receiving the diphtheria, tetanus, and acellular pertussis vaccine; the measles, mumps, and rubella vaccine; and the varicella (chickenpox) vaccine fell to about 94%, just below the 95% mark considered necessary to maintain populationwide immunity. The agency blamed missed checkups because of the pandemic.

Ukraine's fallen include scientists

CONFLICT | As the war in Ukraine grinds into a third month, deaths are mounting—and scholars young and old are among the casualties. Russian soldiers have gunned down scientists in cold blood, including Vasyl Kladko, an x-ray crystallo-grapher at the V.E. Lashkaryov Institute of Semiconductor Physics, slain last month in a Kyiv suburb. In Kharkiv, scientists have died in rocket attacks. And in early April, Andriy Kravchenko, a chemist at the Chuiko Institute of Surface Chemistry, died when a land mine shredded his car near Kyiv. A few days earlier, his team delivered to a hospital the first batch of something he'd spent years developing for battlefield use: a topical coagulant that stanches bleeding until a doctor can reach an injured soldier. "He dreamed it would appear in the first-aid kit of every Ukrainian soldier," says Chuiko colleague Mariia Galaburda. "Such a heavy and painful loss."

Shelled, shot, bombed

This partial list of Ukrainian researchers killed, based on *Science* reporting, shows the war's toll on civilians.

NAME	SPECIALTY	CAUSE OF DEATH
Oleksandr Korsun	Chemistry	Shelling
Oleg Amosov	Economics	Shelling
Yulia Zdanovska	Math	Shelling
Yevhen Khrykov	Education research	Shot
Vasyl Kladko	Physics	Shot
Andriy Kravchenko	Chemistry	Land mine

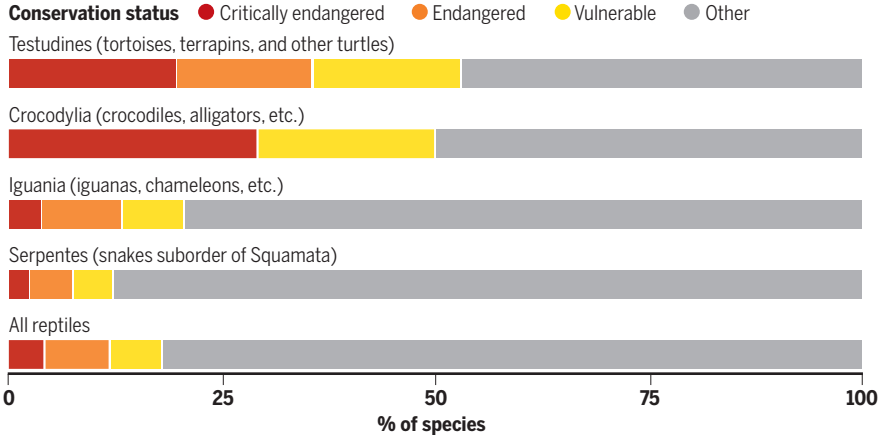
CONSERVATION

Many reptile species face risk of extinction

Half of turtle and crocodile species are in danger, says an extensive global review of reptiles. Experts evaluated data for 10,196 species in 10 major taxonomic groups of reptiles and assessed their conservation status. Hunting and fishing contributed to the high figure for turtle species, the research team reports this week in *Nature*. Overall, 18% of reptile species are threatened; the figure rose to 21% when the scientists extrapolated to include data-deficient species. Reptiles are less at risk than amphibians (41%) and mammals (25%) but worse off than birds (14%).

Grim reckoning

In some orders of reptiles, more than 50% of species were deemed threatened, grouped in one of three categories of risk used by the International Union for Conservation of Nature. (Some orders are not shown.)



U.K. funder regrets IDing critics

WORKPLACE | The funding agency UK Research and Innovation last week apologized for encouraging a UKRI-linked analytics firm to report critical comments made by academics—a move researchers said was a threat to their academic freedom. Scholars had tweeted criticisms about the Researchfish software platform, which collects data for UKRI on the impact of its research grants; the platform then told them it had shared its concerns about the tweet with the researchers' funders. In its apology statement, UKRI said Researchfish had flagged six tweets over 4 years. The agency called its approach "wrong" and promised to stop it. "At no point was this ever intended, or used, to affect current or future grants from UKRI."

Swedish pandemic leader stays put

NEWSMAKERS | Anders Tegnell, former chief epidemiologist at Sweden's Public Health Agency who drew controversy over his country's subdued response to COVID-19 (*Science*, 9 October 2020, p. 159), will remain with the agency after a reported new position consulting for the World Health Organization on pandemic vaccinations didn't materialize. "An agreement could not be reached," the Swedish agency said last week, adding that Tegnell will now work on international issues.

Prenatal tests get FDA warning

GENETICS | Screening evaluations that analyze blood from pregnant people to look for abnormalities in fetal DNA can be unreliable, the U.S. Food and Drug Administration (FDA) warned last week. The tests, marketed in the United States by several companies, are less invasive than diagnostic tests that use amniotic fluid or placenta samples, which in rare cases cause miscarriage. But the blood tests have often incorrectly indicated rare genetic conditions, such as those caused by small missing pieces of chromosomes, FDA said on 19 April. It noted, for example, that follow-up tests fail to confirm seven of 10 positive results for a rare genetic condition called DiGeorge syndrome. Nevertheless, some parents have chosen to end pregnancies based on these results alone, it said. FDA doesn't regulate the tests, and most makers of microdeletion tests don't publicize usage, *The New York Times* reported in January. FDA recommends patients consult a health care provider before deciding to take prenatal blood tests and to confirm any positive result with further tests.

SCIENCE.ORG/NEWS
Read more news from *Science* online.

CREDITS: (GRAPHIC) C. BICKEL/*SCIENCE*; (DATA) N. COX *ET AL.*, *NATURE* 10.1038/S41586-022-04664-7 (2022)

Despite treatment, this bald eagle in Wisconsin became one of many wild birds killed by avian flu.

PHOTO: WISCONSIN HUMANE SOCIETY

INFECTIOUS DISEASE

Deadly flu spreads through North American birds

As largest ever H5N1 outbreak hits poultry and wild species, researchers wonder whether virus is here to stay

By **Erik Stokstad**

When black vultures began to die at Florida's Hontoon Island State Park in February, rangers called in investigators from the state's Fish and Wildlife Conservation Commission. They soon concluded a virus that has devastated domesticated birds worldwide had reached the vultures: a strain of highly pathogenic avian influenza (HPAI) known as H5N1. The vultures had likely acquired the virus from eating infected waterbirds—as well as by cannibalizing their own kind.

Workers removed more than 200 carcasses in a bid to contain the outbreak. But Mark Cunningham, a wildlife veterinarian with the commission, thinks the effort was probably futile. "It's hard to see this chain of infection really breaking anytime soon," he says.

That's a fear shared by researchers and poultry farmers across North America, who in recent weeks have been urgently documenting and trying to contain the continent's largest outbreak of HPAI (p. 459). Since the virus was first spotted in eastern Canada in November 2021, it has been spreading across the continent with migrating waterfowl (see map, p. 442). Poultry farmers have killed nearly 33 million chickens and turkeys in a bid to save other flocks and curb economic losses. Meanwhile, the virus has killed an untold number of wild birds; researchers have so far documented infections in 51 species, including bald eagles and great horned owls. That's more than twice the number of species known to have been infected during the last North American HPAI outbreak, in 2014–15.

HPAI can be far deadlier to birds than seasonal flus are to people, and each outbreak stirs fears about human infection. The current wave has produced no known human cases in North America, however, much to the relief of public health experts already battling COVID-19.

Still, the scope of the HPAI outbreak "boggles the mind," says disease ecologist Nichola Hill of the University of Massachusetts, Boston. She's one of many researchers scrambling to understand how the virus might spread to mammals and whether it will hang on indefinitely in North America, as it has in Europe and Asia. "It's everyone on board, at max capacity," says Susan Shriner, an ornithologist at the U.S. Department of Agriculture, which is helping coordinate the research effort.

The most important HPAI lineage, part of the H5 group of viruses, arose in the late 1990s in domestic geese in Asia. Soon it reassorted with flu strains found in wild waterbirds. In poultry, infections cause pneumonia, seizures, and hemorrhaging with mortality rates of up to 100%. Further mutations enabled those early waves of H5 viruses to infect people—they have killed more than 456 since 2013—raising fears that the viruses could cause a pandemic. But so far, they have not gained the ability to readily spread from person to person.

The H5 viruses did, however, cause catastrophic losses of poultry in Southeast Asia. And migratory birds carried the H5N1 strain out of Asia, first to Europe, where it killed an array of water birds, predatory birds, and scavengers such as buzzards. During the earliest outbreaks, the risk was highest during peak fall migration, when waterfowl arrived in Europe. But in the past 2 years, the virus has become endemic in Europe, present at some level year-round in wild birds. The virus "is not something that is going to go away anytime soon," says Arjan Stegeman, a veterinary epidemiologist at Utrecht University.

Because of the persistence of the virus—and the emergence of an apparently more pathogenic strain of H5N1—Europe has been experiencing ever-worsening HPAI outbreaks in both domestic and wild flocks. Farmers have had to undertake massive culls, and producers of free-range poultry have been forced to move their flocks indoors. Sixty-two wild species have been found infected in Europe and the Middle East in the past 4 months, with some—including barnacle geese, Dalmatian pelicans, and common cranes in Israel—suffering worrisome losses.

In North America, officials have been keeping a wary eye on H5N1. In 2014, migrating birds brought a related virus, H5N8, from Asia to the U.S Pacific Northwest, sparking an outbreak that ultimately caused U.S. farmers in 15 states to kill some 50 million chickens and turkeys and tally $3 billion in losses. This time, H5N1 appears to have arrived from Europe. Last year, after surveys found the strain circulating at high levels among wild birds in Western Europe, U.S.

officials increased their vigilance and requested funds to sample more waterbirds killed by hunters along the Atlantic and Pacific flyways.

In December 2021, several hundred birds died at a petting farm in Newfoundland and tested positive. The next month tests showed a duck killed by a hunter in South Carolina was carrying H5N1. By February, the virus had reached the confluence of the Ohio and Mississippi rivers, where the first farm reported an infection. Since then, migratory birds have spread the virus into the Missouri River Basin and the Great Plains.

Researchers have not yet tested the transmissibility of the virus, but they suspect it spreads more easily than previous strains. That would mean a higher proportion of migratory birds get infected, the geographic spread is wider, and there's a higher prevalence in waterbirds—and hence more spillover into poultry and wild birds, says Bryan Richards, emerging disease coordinator at the U.S. Geological Survey's National Wildlife Health Center.

Genetic analyses of the virus suggest introductions to farms are coming primarily from nearby wild birds. In contrast, researchers believe that during the 2014–15 outbreak of H5N8 humans often accidentally moved the virus from farm to farm.

To prevent infections, many zoos have moved their captive birds indoors or away from visitors. It's much harder to protect wild birds, however, raising fears that the virus could threaten endangered species, especially those with small populations. So far, however, the number of detected infections in wild populations is relatively low, so researchers are "cautiously optimistic ... that we will not see tremendous impacts," Richards says.

Just two songbird species have tested positive: blue jays and crows. At greater risk are water birds, especially those that form dense nesting colonies, and birds that prey on them. Bald eagles often hunt ducks and geese, and some have apparently infected their young by feeding them virus-laden prey, says Rebecca Poulson, a wildlife disease researcher with the Southeastern Cooperative Wildlife Disease Study at the University of Georgia. As of 26 April, 48 eagles known to be infected had died. The eagle deaths—and those of less charismatic birds—are "really heartbreaking," she says. "We've just had our heads hung low some days."

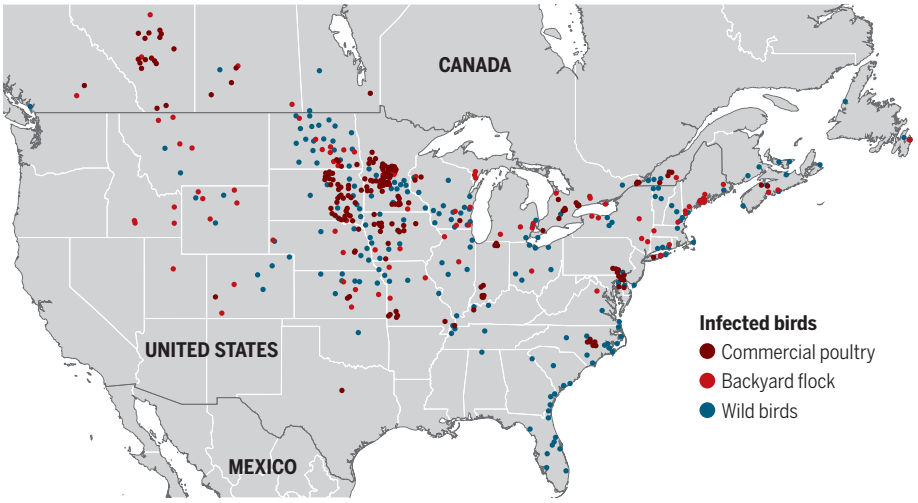
In 2015, Richards says, infections among wild birds petered out in summer as they moved north and dispersed across their nesting grounds, and as lakes and wetlands warmed, creating conditions inhospitable for the virus. This summer, "Knock on wood, we should see a substantial waning of impacts in backyard flocks and commercial operations as well," he says.

But the threat could re-emerge when birds start to migrate south in September. To help farmers stay on guard, the United States will nearly double surveillance efforts. One fear, says Thijs Kuiken, an avian influenza expert at Erasmus University Rotterdam, is that H5N1 will spread south of the U.S. border. Farms there are likely to be more vulnerable.

"People in Central America and South America really need to be aware," he says, "that this virus is likely to arrive on their doorstep—if it hasn't already." ■

A fowl plague

H5N1, a highly pathogenic avian flu virus, arrived from Europe in late 2021. Waterfowl have since spread the virus to other species, including poultry, forcing farmers to kill millions of chickens and turkeys.



CREDITS: (GRAPHIC) K. FRANKLIN/*SCIENCE*; (DATA) PUBLIC DATA COLLATED BY USGS NATIONAL WILDLIFE HEALTH CENTER

COVID-19

India's speedy vaccine approvals come under fire

Critics say regulatory agency lacks key capabilities and independence

By **Priyanka Pulla**

A COVID-19 vaccine named Corbevax looked like a triumph for India's burgeoning drug industry. Because its U.S. developers hadn't claimed a patent on it, an Indian manufacturer named Biological E was able to sell the two-dose protein-based vaccine to the government at the extraordinarily low price of 145 rupees ($1.90) per dose. In March, the country began to give the shots to 12- to 14-year-olds, a group for which India did not yet have a licensed COVID-19 vaccine.

But the celebration was quickly drowned out by questions over whether India's drug regulator, the Central Drugs Standard Control Organization (CDSCO), had properly vetted the vaccine.

In February, CDSCO had authorized the use of Corbevax for adolescents ages 12 to 18. But within weeks, the Indian media outlet *The Wire Science* revealed that the National Technical Advisory Group on Immunisation (NTAGI), an expert group that advises the health ministry on which vaccines to add to the national immunization program, had questioned whether Biological E had shown the vaccine is effective. In adolescents, who are at a lower risk of severe COVID-19, the benefits of a vaccine should be beyond any doubt, NTAGI member Jayaprakash Muliyil tells *Science*: "Anytime you vaccinate children, you have to be extremely careful."

Other CDSCO approvals of COVID-19 vaccines have raised questions as well, both from NTAGI and independent experts. The agency has used "suboptimal" standards on several occasions, says Vineeta Bal, an immunologist at India's National Institute of Immunology. That has led some scientists to ask whether the agency has the capabilities—and is independent enough—to oversee the quality of medicines for India's 1.4 billion people. The implications go beyond India, because

A girl in Kolkata receives a dose of Corbevax, which India began to use in 12- to 14-year-olds in March.

PHOTO: SUKHOMOY SEN/EYEPIX GROUP/FUTURE PUBLISHING/GETTY IMAGES

the country is a major global medicine supplier. The World Health Organization has "prequalified" 54 vaccines produced in India for use elsewhere, and WHO relies on CDSCO to oversee the manufacturers.

CDSCO didn't respond to questions from *Science* about the criticism. In May 2020, India's health ministry appointed a committee to advise it on how to restructure India's drug regulatory system in line with global best practices, but that committee's recommendations haven't been published. It's unclear whether they will address vaccine regulation.

CDSCO has a decent reputation: Based on an extensive assessment, WHO concluded in 2017 that it was a "functional" drug regulator, a distinction only 30% of its counterparts around the world enjoy. (It ranks a step below 11 agencies WHO credits with a "high level of performance," however, including the U.S. Food and Drug Administration and the European Medicines Agency.)

But the pandemic has been challenging for CDSCO, says Gagandeep Kang, a public health microbiologist at Christian Medical College, Vellore. For the first time, Indian manufacturers developed new vaccines rather than me-too versions of shots approved elsewhere, and the agency lacked some types of expertise, Kang says, such as the ability to reanalyze data from key studies. Others suspect CDSCO's decisions have been skewed by political pressure to quickly approve made-in-India vaccines, which have been a point of pride for the ruling Bharatiya Janata Party. "The government has a keenness of publicity, and that has certainly created pressure," says T. Sundararaman, an independent public health analyst based in Puducherry.

In January 2021, for example, the agency greenlit Covaxin, an inactivated-virus vaccine produced by Bharat Biotech, without data from large-scale efficacy trials—only phase 2 data about the immune response generated by the vaccine. By the time the company published data showing 78% efficacy against symptomatic COVID-19, 6 months later, millions of Indians had already received the shot.

NTAGI also differed with CDSCO's assessment when the regulator approved a COVID-19 vaccine named ZyCoV-D for use in both adults and adolescents in August 2021. Produced by Zydus Cadila in Gujarat state, ZyCoV-D is the first DNA vaccine approved by any country for use in humans. CDSCO based its decision on results of a trial in about 28,000 participants over 12 years of age, which found the vaccine 67% efficacious at preventing symptomatic COVID-19.

NTAGI, which does not typically make its advice public, opposed the use in adolescents, Muliyil says; it felt a completely new vaccine platform should only be used in adults at first. Moreover, the phase 3 trial had a single efficacy estimate for all ages, says another NTAGI member who asked not to be identified, even though efficacy can differ by age group. (In Western countries, COVID-19 vaccines for adolescents and children were authorized after separate trials in those age groups.)

In March, CDSCO's reputation took another hit when a WHO inspection of the Covaxin manufacturing facility in Hyderabad found quality control deficiencies, whose nature WHO has not disclosed. WHO recommended that member countries stop using the vaccine, and Bharat voluntarily halted exports. But the company has downplayed the problems and says it will keep selling Covaxin in India. CDSCO did not respond to questions from *Science* about the problems or why it failed to spot them. "It concerns me that CDSCO, the custodian of public health as India's national drug regulator, haven't issued any statements yet on this issue," says Jayanthi Vuppala, an independent expert on good manufacturing practice based in Hyderabad.

Last month's approval of Corbevax—which by now has been given to 30 million adolescents—raised more questions. CDSCO authorized the vaccine for 12- to 18-year-olds based on interim data from a 312-participant study that showed the vaccine triggered a rise in neutralizing antibodies. But NTAGI wasn't convinced the vaccine was entirely responsible for the rise, Muliyil says. Data from the unvaccinated placebo group could have shed light on whether COVID-19 infections were also contributing, but as a preprint posted on 26 April shows, the trial did not assess antibodies in the placebo group. Biological E did not respond to a question about the data.

Since then Botswana has also licensed Corbevax. And on 21 April, an expert committee advising CDSCO recommended the shot's approval for 5- to 12-year-olds in India.

Critics applaud the fact that Corbevax is free from patents, but they worry its approval signals a double standard. "[Would] this vaccine be acceptable for pediatric populations in a high-income country with the data we currently have in the public domain?" asks Boghuma Kabisen Titanji, an infectious disease specialist at Emory University.

Corbevax co-developer Peter Hotez, who leads the Texas Children's Hospital Center for Vaccine Development, says Indian companies have a strong track record for producing WHO-prequalified vaccines. "I'm not aware they apply lower bars and standards," Hotez says. "It would be unfair to think so and in fact it reflects colonial attitudes."

But the controversy has led some Indian pediatricians to not recommend the vaccine. And hasty vaccine approvals, especially for children and adolescents, pose a risk that goes beyond the individual vaccine, Titanji cautions: "It sets a dangerous precedent which could be very damaging for vaccine confidence and uptake down the road." ■

Priyanka Pulla is a science journalist in Bengaluru, India.

SCIENTIFIC COMMUNITY

NYU may hire biologist pushed out of MIT

Med school is in talks with David Sabatini, despite sexual misconduct findings

By **Meredith Wadman**

The New York University (NYU) Grossman School of Medicine is in discussions with biologist David Sabatini about hiring him as a faculty member, according to multiple sources at the school. Since August 2021, Sabatini has been forced out of or fired from three leading institutions for sexual harassment or for violating workplace or consensual sexual relationship policies. The hiring discussions have been taking shape over several weeks, according to NYU sources.

"Any decision about a potential role [for Sabatini] at NYU Grossman School of Medicine would be subject to careful and extensive due diligence and consultation with a broad group of stake holders," the school said in a statement this week.

"David defers to NYU on this process," a Sabatini spokesperson said.

In an email to faculty and staff last week, the medical school's dean decried what he called "cancel culture." And last month, a prominent philanthropist defended Sabatini to a crowd of elite researchers in New York City.

But numerous faculty members at NYU's medical school are disturbed by the prospect of hiring Sabatini, saying the move sends a message of tolerance for sexual misconduct and hostile lab environments. "We are deeply concerned that recruiting an individual found guilty of having broken the policies of prestigious institutions such as MIT [the Massachusetts Institute of Technology], the Whitehead Institute and HHMI [the Howard Hughes Medical Institute] may profoundly damage our culture and reputation," reads a letter to the dean and vice dean that had 66 faculty signatures as of midafternoon Tuesday.

"I'm concerned for the culture of the institution," said one faculty member, who asked to be anonymous for fear of retaliation by administrators. "The accusations made are really serious. … Is this signaling that in the future students that come forward for any kind of faculty misconduct will be disregarded?"

However, some current and former Sabatini lab members are enlisting anonymous signatures for an online letter supporting him. The letter says the signers "never experienced or observed an abusive lab culture or a sexualized lab environment, and we did not witness sexual harassment. … Working in David's lab was one of the most influential and fortunate events of our careers. … We hope that this letter will contribute positively in your decision to hire David."

Sabatini, a prominent cancer biologist, resigned from MIT earlier this month after senior officials recommended revoking his tenure. They found he had violated MIT's policy on consensual sexual relationships and expressed "significant concerns regarding his professional behavior to some lab members," President L. Rafael Reif wrote to MIT faculty on 1 April.

Cancer biologist David Sabatini, pictured at a 2019 awards dinner.

PHOTO: SYLVAIN GABOURY/PATRICK MCMULLAN VIA GETTY IMAGES

In August 2021, Sabatini was forced out of the Whitehead. An outside law firm had "found that Dr. Sabatini violated the Institute's policies on sexual harassment, among other Whitehead policies unrelated to research misconduct," Director Ruth Lehmann wrote to Whitehead staff. HHMI, which had funded Sabatini's Whitehead lab, fired him at the same time, saying he had violated workplace behavior policies.

The institutions have not provided details about the behaviors that led to Sabatini's removal, but the Whitehead compiled a 229-page report, according to a defamation lawsuit Sabatini filed in October 2021 against the Whitehead, Lehmann, and a junior colleague who told investigators that Sabatini had sexually harassed her. He had served on her Ph.D. thesis committee, and she became a Whitehead fellow with her own lab in a program he was soon appointed to direct. She filed a counterclaim against Sabatini in December, arguing his lawsuit was frivolous and retaliatory. (*Science* does not name targets of alleged sexual misconduct without their permission.)

Sabatini's hire may have support from a benefactor who could make up some of the funding Sabatini lost when HHMI fired him. Billionaire hedge fund founder Bill Ackman of Pershing Square Capital Management supports coveted awards for early-career cancer researchers. Sabatini is a reviewer for those awards and attended a 1 March Pershing Square Foundation dinner at Manhattan's Le Bernardin restaurant. In his remarks, Ackman bemoaned what he called Sabatini's unfair treatment and said sidelining Sabatini's work would be detrimental to the field. (Ackman was not available for comment.)

On 21 April, NYU medical school dean and CEO Robert Grossman sent his monthly email to medical school faculty and staff, titling it "Civility Rules." "A mob feels compelled to stridently 'cancel' someone with different thoughts, or to baselessly attack an individual in ways that can be difficult to disprove," he wrote. "[This is] tolerated in academia, where nameless accusers can disparage a colleague's science and life's work, even when their claims are unfounded." Grossman declined to comment further.

Former Sabatini postdoc Anne Carpenter, now a computational biologist at the Broad Institute, noted on Twitter that in early April she was asked to sign the anonymous letter of support. "I will not sign," Carpenter wrote in a tweet thread on 10 April. She wrote that although the MIT investigation did not involve her, she was not surprised by its findings.

In his lawsuit, Sabatini had cited Carpenter's successful career as an example that "contradicts a finding that women were disadvantaged" in his lab. But Carpenter called that logic "appalling" in her thread. "My success is not evidence that he did not break policies," she wrote. "It does not prove that he treated men and women equally in the lab." ■

At the Polytechnic University of Nicaragua in 2018, a memorial honors students killed in protests.

PHOTO: INTI OCON/AFP/GETTY IMAGES

SCIENCE AND POLITICS

Nicaragua's universities stagger under government pressure

Attacks on higher education autonomy put research—and researchers—at risk

By **Sofia Moutinho**

Repression of academia in Nicaragua has escalated over the past 3 months. President Daniel Ortega's government has closed eight private universities and confiscated their assets, effectively ended university autonomy, and erected barriers to foreign research collaborations. "They are killing the universities," says chemist Ernesto Medina, former dean of American University, a private institution in Managua that remains open. "All these measures serve to silence the critical voices in academia and suffocate critical thinking."

Ortega has been president for 15 years, having won reelections condemned internationally as farcical. In April 2018, students took to the streets in antigovernment protests. The government responded with unprecedented repression (*Science*, 21 December 2018, p. 1338). Police and paramilitary groups killed more than 300 people and arrested more than 1000, according to Amnesty International. After writing two open letters asking Ortega to stop the "irrational violence," Medina was forced out of his university position. He fled to Germany in July 2020, becoming one of more than 200,000 Nicaraguans in exile. "It is not safe to be there," he says.

Since then, conditions for Nicaragua's small scientific community have worsened. After the 2018 protests, international research conferences were canceled and funding for scientific institutions, such as the Nicaraguan Academy of Sciences, was cut.

Researchers say they must be particularly cautious about any work that could be interpreted as political. Research on democracy, gender, and human rights has been especially hard hit. A social scientist says his group can no longer conduct surveys and now publishes under pseudonyms if mentioning government abuses. (Several researchers who commented for this article requested anonymity, fearing retaliation.) Research about COVID-19 in Nicaragua also faces obstacles, as the government has understated case numbers and fired doctors and public health experts who express concern about the crisis.

International collaborations, which had helped sustain science in Nicaragua, are withering. In February, the government canceled the permits of foreign institutions running educational and research programs, such as Florida International University and Michigan State University. A law passed in September 2020 requires any Nicaraguan working with international organizations to self-identify as a "foreign agent." But foreign agents are considered traitors, researchers say. "It practically destroyed any possibility of international collaboration," a biologist says. U.S. researchers who collaborate with Nicaraguans declined to comment for this article, fearing reprisal against their colleagues.

The government claims the eight private universities shut down since February had failed to disclose financial details. Sources at one closed institution, the Polytechnic University of Nicaragua, believe the real reason was punishment for students' involvement in the 2018 protests.

The National Council of Universities (CNU), a government body that coordinates national policy on higher education, announced in February that three new public universities would replace the closed ones, offering enrollment to 20,000 former students. But sources say many students have given up on academics, fearing persecution. The new institutions "reflect the reality of Nicaraguan education, submitted to a totalitarian regime that has no interest besides keeping the students under control," Medina says.

Recent law changes also increased the power of CNU, which is now responsible for reviewing academic programs, approving academic hires, and selecting deans in all public universities. "It is the end of universities' autonomy," a former CNU member says.

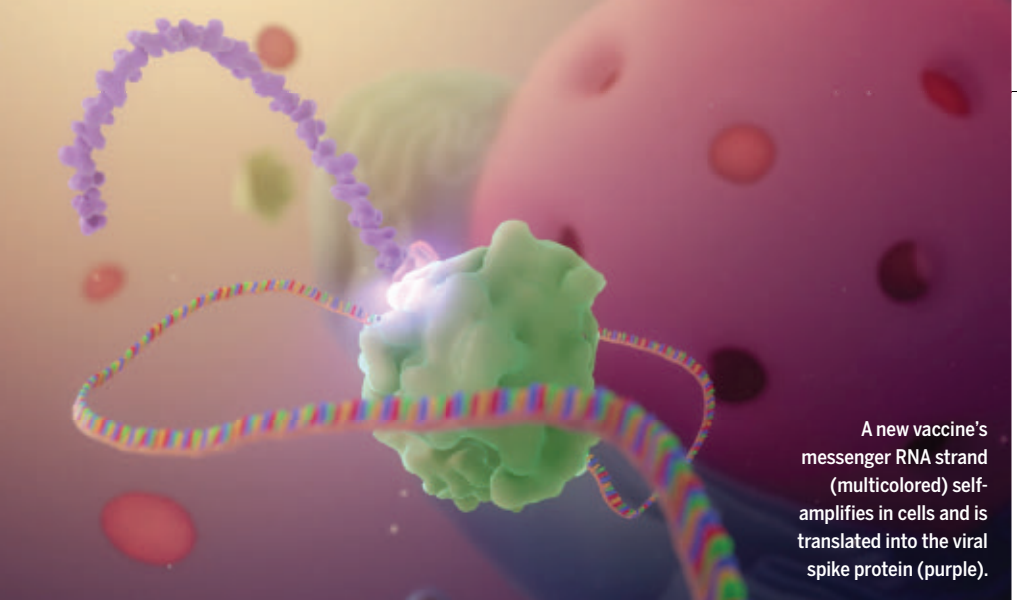
Some of the 40 or so private universities that remain open face economic strangulation. One particular target is the Central American University (UCA), which officially opposed the government's violence toward protesters. This month, Ortega's government canceled a state fund for UCA that provided scholarships for 4000 low-income students. A scientist who is a professor there says the university "is a place of resistance," and faculty will keep teaching until it is forced to close.

"We are in a stage of consolidation of totalitarianism in education," says former UCA law professor María Asunción Moreno, who is now in exile. "It is no exaggeration to say there is no university in Nicaragua anymore." In 2019, Moreno led law students and lawyers who collaborated with the Inter-American Commission on Human Rights to document government abuses. In July 2020, she escaped Nicaragua after 12 days hiding from arrest.

Medina estimates more than 2000 Nicaraguan students and young professionals, including many scientists, have fled, mainly to Costa Rica and Mexico. He hopes universities in other countries will create scholarships for exiled students. "We have to train people so they can think about the changes we'll want when we have democracy again," he says.

Molecular biologist Helena Nader, co-chair of the Inter-American Network of Academies of Sciences, says Nicaragua is an example of the growing threats to science and democracy in Latin America, including in Venezuela and El Salvador. "What is happening in Nicaragua is very serious, and the world is silent," she says. ■

Sofia Moutinho is a journalist based in Rio de Janeiro.



A new vaccine's messenger RNA strand (multicolored) self-amplifies in cells and is translated into the viral spike protein (purple).

IMAGE: ARCTURUS THERAPEUTICS

COVID-19

First self-copying mRNA vaccine proves itself in pandemic trial

Twist on current vaccines reduces dose, eases distribution

By **Jon Cohen**

A third messenger RNA (mRNA) vaccine for COVID-19 appears to have proved its worth. And although it is more than a year behind the Moderna and Pfizer-BioNTech vaccines now seen as gold standards, the new one may come with significant advantages: easier storage, along with lower cost because its "self-amplifying" design allows for smaller doses.

Arcturus Therapeutics of San Diego, which staged a placebo-controlled trial of its candidate in more than 17,000 participants in Vietnam, announced last week in a press release that the vaccine had 55% efficacy against symptomatic COVID-19 and provided 95% efficacy against severe illness and death. "It's a huge accomplishment that for the first time a self-amplifying RNA vaccine has been shown to be safe and effective," says Deborah Fuller, a vaccinologist at the University of Washington School of Medicine who is an adviser to HDT Bio, which has its own self-amplifying COVID-19 mRNA vaccine in human studies.

Arcturus's success may also help make mRNA vaccines more broadly accessible. Its candidate incorporates a freeze-drying process to transform the mRNA-filled solution into a powder that's hydrated before use. As a result it requires far less refrigeration than the conventional, liquid mRNA vaccines. And Vietnam's Vinbiocare Biotechnology, which collaborated with Arcturus and has submitted the trial data to the country's regulators for emergency use authorization, hopes to manufacture the product there.

The Pfizer-BioNTech and Moderna vaccines contain mRNA that codes for the SARS-CoV-2 spike protein. When the vaccines are injected, they deliver the mRNA to cells, which make copies of spike and then clear the foreign genetic material within a few days. Arcturus's self-amplifying vaccine and others in development include enzymes from alphaviruses to repeatedly copy the genetic strand inside a cell, enabling it to persist in the body for more than twice as long.

Some researchers have cautioned that self-amplifying vaccines cannot use an mRNA modification that seemed key: the replacement of the natural RNA building block uridine with pseudouridine. Studies have shown the swap leads to higher levels of the spike protein and lower production of immune chemicals called cytokines that can cause side effects. A conventional mRNA vaccine made by CureVac failed in an efficacy trial last year, and it didn't use pseudouridine. But Arcturus says its efficacy findings show the swap isn't needed. "It's a big deal for the field," says Pad Chivukula, the company's chief scientific officer.

The trial gave participants two doses, each containing 5 micrograms of the self-amplifying mRNA, spaced 28 days apart. The Pfizer-BioNTech and Moderna vaccines use 30-microgram and 100-microgram doses, respectively, for the first two shots.

The 55% efficacy of Arcturus's vaccine against symptomatic infection is below the 90% to 95% seen in trials of the first two mRNA vaccines. But those two encoded a version of spike that closely matched the SARS-CoV-2 variant then circulating. The Arcturus candidate encoded the same spike yet had to protect against the Delta and Omicron variants, which have evolved dramatically from the ancestral strain, decreasing the power of vaccine-triggered antibodies.

Fuller says the current real-world effectiveness of existing mRNA vaccines may be in the same ballpark. Of the 43 severe cases of COVID-19 recorded by Arcturus during the trial, only two were in the vaccinated group, and nine of the 10 people who died of COVID-19 received the placebo.

"These are indeed exciting results," says chemist Benjamin Pierce, who is helping run a Ugandan trial of a self-amplifying mRNA COVID-19 vaccine made by Imperial College London. "The low dose used here—six to 20 times lower than approved RNA vaccines—further indicates that self-amplifying RNA technology has such potential. I look forward to seeing more of the data from the trial."

Fuller says a single dose of the self-amplifying vaccine could be used for the two primary doses, giving it an even clearer benefit over its conventional relatives. A booster later might still be warranted, as is encouraged for the current mRNA vaccines. But self-amplifying mRNAs could lead to more durable immune responses, Fuller suggests.

When the Arcturus trial began, less than 15% of the eligible Vietnamese population had received a single shot of a COVID-19 vaccine. Now, the figure is 80%, which raises the question of how the vaccine will perform in the vast majority of people who have already been vaccinated or naturally exposed to SARS-CoV-2. Arcturus hopes to soon launch a 2400-person trial to assess its value as a booster shot. This trial will aim to show the vaccine boosts antibody responses that other studies have shown correlate with protection—though no new COVID-19 vaccine has yet received authorization from strict U.S. or European regulators based on such "immunobridging" data.

The Pfizer-BioNTech collaboration and Moderna have been criticized for not quickly sharing their manufacturing skills and intellectual property with developing countries, which have had relatively little access to their mRNA vaccines. Arcturus, in contrast, has agreed to a technology transfer deal with Vinbiocare, which is building a vaccine manufacturing plant in Hanoi.

Some observers question whether the Arcturus vaccine is making its debut too late, at least for primary vaccination. Chivukula is confident it will find a market in countries that have far lower vaccination rates than Vietnam and stresses it will be at "a price point that everyone can afford." ■

SCIENCE & SECURITY

Math professor's trial next test of China Initiative

Mingqing Xiao is fighting U.S. charges that he failed to disclose ties to Chinese institutions

By **Jeffrey Mervis**

This week, a federal jury in Benton, Illinois, began hearing the U.S. government's evidence against Mingqing Xiao, a professor of applied math at Southern Illinois University (SIU), Carbondale, accused of failing to disclose his research ties to China while receiving a grant from the National Science Foundation (NSF).

Each day, a small cluster of Xiao's colleagues are attending the proceedings, which opened on 25 April, wearing buttons that proclaim "We stand with Mingqing Xiao." It is the fourth jury trial of a U.S. academic resulting from a controversial law enforcement effort known as the China Initiative. The outcome will be another important test of the U.S. government's hard-edged strategy of prosecuting academic scientists—some two dozen to date—whose actions are alleged to have undermined national security.

The China Initiative was launched in 2018 to prevent the Chinese government from stealing intellectual property, Department of Justice officials said at the time. However, Xiao and most of the other prosecuted scientists have been accused of financial fraud and tax violations, not of improperly transferring U.S.-funded research results. Most of the cases have been settled with plea deals or dropped. Three jury trials have resulted in two convictions and an acquittal.

Xiao's background as a Chinese immigrant who has built an academic career in the United States is typical of many of the defendants. Born and raised in Guangdong, China, Xiao came to the United States after the 1989 Tiananmen Square uprising, ultimately earning his Ph.D. from the University of Illinois, Urbana-Champaign. He joined the SIU faculty in 2000, became a U.S. citizen in 2006, and a year later was promoted to full professor. In 2016, SIU named him an outstanding scholar, and in 2020 the university's public radio station gave him its "neighborly" award for founding and running a free Saturday morning math academy for children.

In presenting its case this week against Xiao, the government was expected to argue that the 60-year-old academic broke the law when he obtained an NSF grant in 2019 to develop new mathematical tools for analyzing high-dimensional data sets. It alleges that Xiao had ties to Shenzhen University and Guangdong University of Technology that he failed to disclose to SIU and NSF, and that he received compensation from those entities that he didn't list on his federal tax returns. Prosecutors will assert that he did so deliberately, with the intent to defraud the government. If convicted, he faces up to 20 years in prison and a sizable fine.

Xiao's lawyer, Ryan Poscablo of Steptoe & Johnson, will argue there's insufficient evidence to prove the allegations and, moreover, that there was no criminal intent behind any incorrect statements Xiao might have made. It is not illegal for U.S. academics to receive research funding from foreign sources. In fact, until recently most U.S. universities encouraged faculty to hook up with a burgeoning scientific powerhouse such as China, and being of Chinese descent was seen as an asset. But any ties must be disclosed when applying for a federal grant. The nature of such disclosures, and whether some of those relationships were deliberately hidden, have emerged as central issues in many China Initiative cases.

Mingqing Xiao went on trial this week in a federal courtroom in Illinois.

Many academics and civil rights groups have vehemently criticized the prosecutions, calling the initiative a witch hunt against U.S. scientists of Chinese origin and a misguided attack on legitimate research collaborations. Two months ago, Assistant Attorney General Matthew Olsen acknowledged such criticisms and rebranded the initiative as a "strategy to counter nation-state threats" (*Science*, 4 March, p. 945).

But opponents say they see few signs the government has changed course. "From where I sit, the demise of the China Initiative is greatly exaggerated," says Peter Zeidenberg, an attorney for Franklin Tao, a University of Kansas, Lawrence, chemical engineer convicted earlier this month in a decision that some legal experts think could be overturned.

Xiao was indicted in April 2021, and the university immediately put him on paid administrative leave and launched its own investigation. Xiao's academic colleagues initially refrained from making any public statements on his behalf. "We didn't have any information beyond what the Justice Department had said, and we didn't want to jeopardize his defense," says Edward Benyas, an SIU music professor who helped organize the daily vigils. Many also feared that, as state employees, they might face repercussions.

However, in November 2021 the SIU Carbondale Faculty Association called on the government to drop the charges and asked the university to reinstate Xiao and finance his legal defense. The university has declined further comment on what a spokesperson called "a personnel matter."

Asian American groups that have spoken out on behalf of other academics targeted by the China Initiative acknowledge they have done little on Xiao's behalf. They say meager resources and a lack of bandwidth have prevented them from being more active.

Although Xiao has continued to receive his university salary, he has been overwhelmed with legal bills and has drained his life savings. A GoFundMe campaign has raised $26,000 toward a goal of $350,000, and supporters say the real need tops $800,000.

Benyas and other Xiao supporters don't expect their presence in the courtroom to have an impact on the jury's deliberations. "But we think it's important to show that people are supporting him," he says. "It's the least we can do." ■

A QUANTUM SENSE FOR DARK MATTER

By harnessing the strange rules of the subatomic realm, quantum sensors could solve one of the universe's biggest mysteries

By **Adrian Cho**

Kent Irwin has a vision: He aims to build a glorified radio that will reveal the nature of dark matter, the invisible stuff that makes up 85% of all matter. For decades, physicists have struggled to figure out what the stuff is, stalking one hypothetical particle after another, only to come up empty. However, if dark matter consists of certain nearly massless particles, then in the right setting it might generate faint, unquenchable radio waves. Irwin, a quantum physicist at Stanford University, plans to tune in to that signal in an experiment called Dark Matter Radio (DM Radio).

No ordinary radio will do. To make the experiment practical, Irwin's team plans to transform it into a quantum sensor—one that exploits the strange rules of quantum mechanics. Quantum sensors are a hot topic, having received $1.275 billion in funding in the 2018 U.S. National Quantum Initiative. Some scientists are employing them as microscopes and gravimeters. But because of the devices' unparalleled sensitivity, Irwin says, "dark matter is a killer app for quantum sensing."

DM Radio is just one of many new efforts to use quantum sensors to hunt the stuff. Some approaches detect the granularity of the subatomic realm, in which matter and energy come in tiny packets called quanta. Others exploit the trade-offs implicit in the famous Heisenberg uncertainty principle. Still others borrow technologies being developed for quantum computing. Physicists don't agree on the definition of a quantum sensor, and none of the concepts is entirely new. "I would argue that quantum sensing has been happening in one form or another for a century," says Peter Abbamonte, a condensed matter physicist and leader of the Center on Quantum Sensing and Quantum Materials at the University of Illinois, Urbana-Champaign (UIUC).

Still, Yonatan Kahn, a theoretical physicist at UIUC, says quantum sensors open the way to testing new ideas for what dark matter might be. "You shouldn't just go blindly looking" for dark matter, Kahn says. "But even if your model is made of bubblegum

CREDITS: (X-RAY) NASA/CXC/CFA/M. MARKEVITCH *ET AL.*; (OPTICAL) NASA/STSCI; MAGELLAN/U. ARIZONA/D. CLOWE *ET AL.*; (LENSING MAP) NASA/STSCI; ESO WFI; MAGELLAN/U. ARIZONA/D. CLOWE *ET AL.*

A collision of galaxy clusters separated gas (pink) from dark matter (blue), mapped from subtle gravitational distortions in the images of background galaxies.

and paperclips, if it satisfies all cosmological constraints, it's fair game." Quantum sensing is essential for testing many of those models, Irwin says. "It can make it possible to do an experiment in 3 years that would otherwise take thousands of years."

ASTROPHYSICAL EVIDENCE for dark matter has accreted for decades. For example, the stars in spiral galaxies appear to whirl so fast that their own gravity shouldn't keep them from flying into space. The observation implies that the stars circulate within a vast cloud of dark matter that provides the additional gravity needed to rein them in. Physicists assume it consists of swarms of some as-yet-unknown fundamental particle.

In the 1980s, theorists hypothesized what soon became the leading contender: weakly interacting massive particles (WIMPs). Emerging in the hot soup of particles after the big bang, WIMPs would interact with ordinary matter only through gravity and the weak nuclear force, which produces a kind of radioactive decay. Like the particles that convey the weak force, the *W* and *Z* bosons, WIMPs would weigh roughly 100 times as much as a proton. And just enough WIMPs would naturally linger—a few thousand per cubic meter near Earth—to account for dark matter.

Occasionally a WIMP should crash into an atomic nucleus and blast it out of its atom. So, to spot WIMPs, experimenters need only look for recoiling nuclei in detectors built deep underground to protect them from extraneous radiation. But no signs of WIMPs have appeared, even as detectors have grown bigger and more sensitive. Fifteen years ago, WIMP detectors weighed kilograms; now, the biggest contain several tons of frigid liquid xenon.

"It can make it possible to do an experiment in 3 years that would otherwise take thousands of years."
Kent Irwin, Stanford University

The second most popular candidate—and one DM Radio targets—is the axion. Far lighter than WIMPs, axions are predicted by a theory that explains a certain symmetry of the strong nuclear force, which binds quarks into trios to make protons and neutrons. Axions would also emerge in the early universe, and theorists originally estimated they could account for dark matter if the axion has a mass between one-quadrillionth and 100-quadrillionths of a proton.

In a strong magnetic field, an axion should sometimes turn into a radio photon whose frequency depends on the axion's mass. To amplify the faint signal, physicists place in the field an ultracold cylindrical metal cavity designed to resonate with radio waves just as an organ pipe rings with sound. The Axion Dark Matter Experiment (ADMX) at the University of Washington, Seattle, scans the low end of the mass range, and an experiment called the Haloscope at Yale Sensitive to Axion CDM (HAYSTAC) at Yale University probes the high end. But no axions have shown up yet.

In recent years physicists have begun to consider other possibilities. Maybe axions are either more or less massive than previously estimated. Instead of one type of particle, dark matter might even consist of a hidden "dark sector" of multiple new particles that would interact through gravity but not the three other forces, electromagnetism and the weak and strong nuclear forces. Rather, they would have their own forces, says Kathryn Zurek, a theorist at the California Institute of Technology. So, just as photons convey the electromagnetic force, dark photons might convey a dark electromagnetic force. Dark and ordinary electromagnetism might intertwine so that rarely, a dark photon could morph into an ordinary one.

To spot such quarry, dark matter hunters have turned to quantum sensors—a shift partly inspired by another hot field: quantum computing. A quantum computer flips quantum bits, or qubits, that can be set to 0, 1, or, thanks to the odd rules of quantum mechanics, 0 and 1 at the same time. That may seem irrelevant to hunting dark matter, but such qubits must be carefully controlled and shielded from external interference, exactly what dark matter hunters already do with their detectors, says Aaron Chou, a physicist at Fermi National Accelerator Laboratory (Fermilab) who works on ADMX. "We have to keep these devices very, very well isolated from the environment so that when we see the very, very rare event, we're more confident that it might be due to the dark matter."

The interest in quantum sensors also reflects the tinkerer culture of dark matter hunters, says Reina Maruyama, a nuclear and particle physicist at Yale and co-leader of HAYSTAC. The field has long attracted people interested in developing new detectors and in quick, small-scale experiments, she says. "This kind of footloose approach has always been possible in the dark matter field."

FOR SOME NOVEL SEARCHES, the simplest definition of a quantum sensor may do: It's any device capable of detecting a single quantum particle, such as a photon or an energetic electron. "I call a quantum sensor something that can detect single quanta in whatever form that takes," Zurek says. That's what is needed for hunting particles slightly lighter than WIMPs and plumbing the dark sector, she says.

Such runty particles wouldn't produce detectable nuclear recoils. A wispy dark sector particle could interact with ordinary matter by emitting a dark photon that morphs into an ordinary photon. But that low-energy photon would barely nudge a nucleus.

In the right semiconductor, however, the same photon could excite an electron and enable it to flow through the material. Kahn and Abbamonte are working on an extremely sensitive photodiode, a device that produces an electrical signal when it absorbs light. Were such a device shielded from light and other forms of radiation and cooled to near absolute zero to reduce noise, a dark matter signal would stand out as a

steady pitter-pat of tiny electrical pulses.

The trick is to find a semiconductor sensitive to very low-energy photons, Kahn says. The industrial standard, silicon, releases an electron when it absorbs a photon with an energy of at least 1.1 electron volts (eV). To detect dark sector particles with masses as low as 1/100,000th that of a proton, the material would need to unleash an electron when pinged by a photon of just 0.03 eV. So Kahn, Abbamonte, and colleagues at Los Alamos National Laboratory are exploring "narrow bandgap" semiconductors such as a compound of europium, indium, and antimony.

Even lighter dark-sector particles would create photons with too little energy to liberate an electron in the most sensitive semiconductor. To hunt for them, Zurek and Matt Pyle, a detector physicist at the University of California, Berkeley, are developing a detector that would sense the infinitesimal quantized vibrations set off when a dark photon creates an ordinary photon that pings a nucleus. It would "only rattle that nucleus and produce a bunch of vibrations," Pyle says. "So the detectors must be fundamentally different."

Their detector consists of a single crystal of material composed of two types of ions with opposite charges, such as gallium arsenide. The feeble photon spawned by a dark photon would nudge the different ions in opposite directions, setting off quantized vibrations called optical phonons. To detect these vibrations, Zurek and Pyle dot the crystal with small patches of tungsten and chill it to temperatures near absolute zero, where tungsten becomes a superconductor that carries electricity without resistance. Any phonons would slightly warm the tungsten, reducing its superconductivity and leading to a noticeable spike in its resistance.

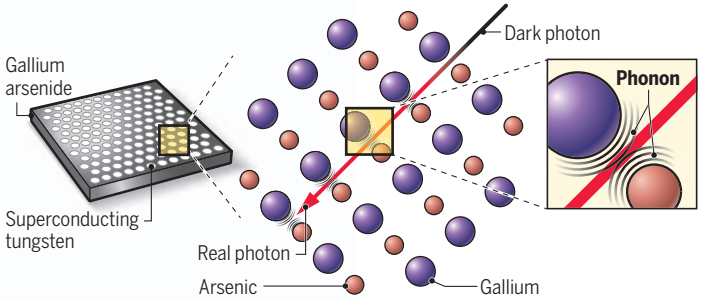
Within 5 years, the researchers hope to improve their detector's sensitivity by a factor of 10 so that they can sense a single phonon and hunt dark-sector particles weighing one-millionth as much as a proton. To provide the dark matter, such particles would have to be so numerous that a detector weighing just a few kilograms should be able to spot them or rule them out. And because so few experiments have probed this mass range, even little prototype detectors unshielded from background radiation can yield interesting data, Pyle says. "We run just in our lab aboveground, and we can get world-leading results."

Particles and waves

Quantum detectors include devices that can detect a single quantum, such as a photon, and devices that exploit a quantum trade-off to measure one variable more precisely at the cost of greater uncertainty in another.

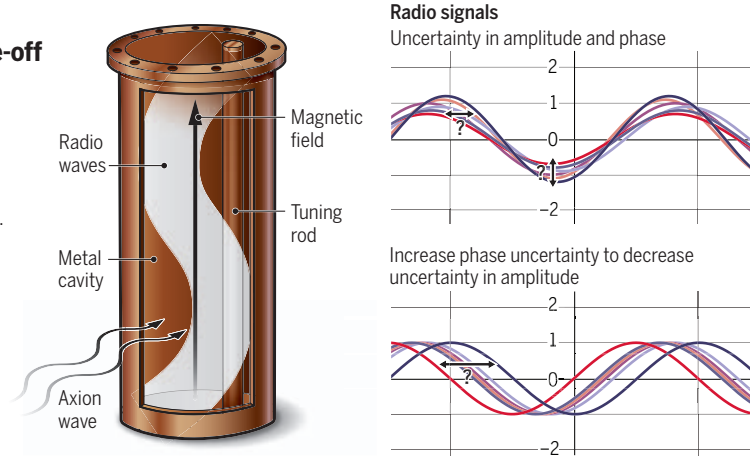
Just one click

A dark matter candidate called a dark photon could morph into an ordinary photon that would trigger a quantized vibration in a crystal. The vibration, or phonon, would warm superconducting heat sensors on the crystal.



Quantum trade-off

Within a resonating cavity, a wave of hypothetical axions could transform into faint radio waves, uncertain in both amplitude and phase. Quantum techniques could reduce the uncertainty in the amplitude while increasing that in the wave's irrelevant phase.



SOME PHYSICISTS ARGUE that true quantum sensors should do something more subtle. The Heisenberg uncertainty principle states that if you simultaneously measure the position and momentum of an electron, the product of the uncertainties in those measurements must exceed a "standard quantum limit." That means no measurement can yield a perfectly precise result, no matter how it's done. However, the principle also implies you can swap greater uncertainty in one measurement for greater precision in the other. To some physicists, a quantum sensor is one that exploits that trade-off.

Physicists are using such schemes to enhance axion searches. To make up dark matter, those lightweight particles would be so numerous that en masse they'd act like a wave, just as sunlight acts more like a light wave than a hail of photons. So with their metal cavities, ADMX and HAYSTAC researchers are searching for the conversion of an invisible axion wave into a detectable radio wave.

Like any wave, the radio wave will have an amplitude that reveals how strong it is and a phase that marks its exact synchronization relative to whatever ultraprecise clock you might choose. Conventional radio circuits measure both and run into a limit set by the uncertainty principle. But axion hunters care only about the signal's amplitude—is a wave there or not?—and quantum mechanics lets them measure it with greater precision in exchange for more uncertainty in the phase.

HAYSTAC experimenters exploit that trade-off to tamp down noise in their experiment. The vacuum—the backdrop for the measurement—can itself be considered a wave. Although that vacuum wave has on average zero amplitude, its amplitude is still uncertain and fluctuates to create noise. In HAYSTAC a special amplifier reduces the vacuum's amplitude fluctuations while allowing those in the irrelevant phase to grow bigger, causing any axion signal to stand out more readily. Last year, HAYSTAC researchers reported in *Nature* that they had searched for and ruled out axions in a narrow range around 19-quadrillionths of a proton mass. By squeezing the noise, they increased the speed of the search by a factor of 2, Maruyama says, and validated the principle.

Such "squeezing" has been demonstrated for decades in laboratory experiments with lasers and optics. Now, Irwin says, "These techniques for beating the standard quantum limit [have] been used to actually do something better, as opposed to do something in a demonstration." In the DM Radio experiment, he hopes to use a related technique to probe for even lighter axions as well as dark photons.

GRAPHIC: V. ALTOUNIAN/*SCIENCE*

Instead of a resonating cavity, DM Radio consists of a radio circuit containing a charge-storing capacitor and a current-storing inductor—a carefully designed coil of wire—both placed in a magnetic field. Axions could convert to radio waves within the inductor coil to create a resonating signal in the circuit at a certain frequency. Researchers can also look for dark photons by reconfiguring the coil and turning off the magnetic field.

To read out the signal, Irwin's scheme plays on another implication of quantum mechanics, that by measuring a system's state you may change it. The researchers couple their resonating circuit to a second, higher frequency circuit, so that, much as in AM radio, any dark matter signal would make the amplitude of the higher frequency carrier wave warble. The stronger the coupling, the bigger the warbling, and the more prominent the signal. But stronger coupling also injects noise that could stymie efforts to measure dark matter with greater precision.

Again, a quantum trade-off comes to the rescue. The researchers modify their carrier wave by injecting a tiny warble at the frequency they hope to probe. Just by random chance, that input warble and any dark matter signal will likely be somewhat out of sync, or phase. But the dark matter wave can be thought of as the sum of two components: one that's exactly in sync with the added signal and one that's exactly out of sync with it—much as any direction on a map is a combination of north-south and east-west. The experiment is designed to measure the in-sync component with greater precision while injecting all the disturbance into the out-of-sync component, making the measurement more sensitive and accelerating the rate at which the experiment can scan different frequencies.

Irwin and colleagues have already run a small prototype of the experiment. They are now building a larger version, and ultimately they plan one with a coil that has a volume of 1 cubic meter. Implementing the quantum sensing is essential, Irwin says, as without it, scanning the entire frequency range would take thousands of years.

SOME DARK MATTER HUNTERS are explicitly borrowing hardware from quantum computing. For example, Fermilab's Chou and colleagues have used a superconducting qubit—the same kind Google and IBM use in their quantum computers—to perform a proof-of-principle search for dark photons in a very narrow energy range. Like a smaller version of ADMX or HAYSTAC, their experiment centers on a resonating cavity, this one drilled into the edge of an aluminum plate. There a dark photon could convert into radio waves, although at a higher frequency than in ADMX or HAYSTAC. Ordinarily, experimenters would bleed the radio waves out through a hole in the cavity and measure them with a low-noise amplifier. However, the tiny cavity would generate a signal so faint it would drown in noise from the amplifier itself.

The qubit sidesteps that problem. Like any other qubit, the tiny superconducting circuit can act like a clock, cycling between different combinations of 0 and 1 at a rate that depends on the difference in energy between the circuit's 0 and 1 states. That difference in turn depends on whether there are any radio photons in the cavity. Even one is enough to speed up the clock, Chou says. "We're going to stick this artificial atomic clock in the cavity and see if it still keeps good time."

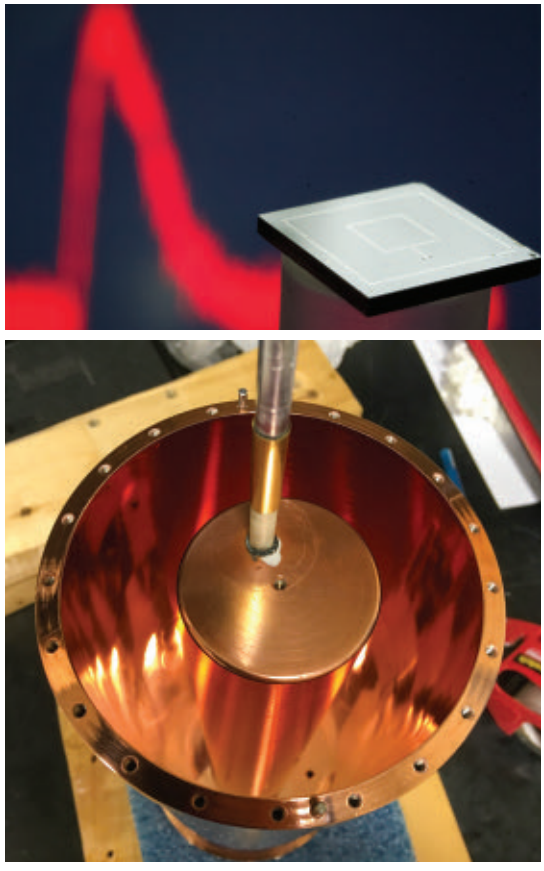
The measurement probes only the amplitude of the radio waves and not their phase, obtaining greater precision in the former in exchange for greater uncertainty in the latter, the team reported last year in *Physical Review Letters*. It might speed up dark photon searches by as much as a factor of 1300, Chou says, and it could be extended to search for axions, if researchers could apply a magnetic field to the cavity while shielding the sensitive qubit.

One group has invented a scheme to search for WIMPs using another candidate qubit: a so-called nitrogen vacancy (NV) center within a diamond crystal. In an NV, a nitrogen atom replaces a carbon atom in the crystal lattice and creates an adjacent, empty site that collects a pair of electrons that can serve as qubit. A WIMP passing through a diamond can bump carbon atoms out of the way, leaving a trail of NVs roughly 100 nanometers long, says Ronald Walsworth, an experimental physicist at the University of Maryland, College Park. The NVs will absorb and emit light of specific wavelengths, so the track can be spotted clearly with fluorescence microscopy.

That scheme has little to do with quantum computing, but it would address a looming problem for WIMP searches. If current liquid xenon detectors get much bigger, they should start to see well-known particles called neutrinos, which stream from the Sun. To tell a WIMP from a neutrino, physicists would need to know where a particle came from, as WIMPs should come from the plane of the Galaxy rather than the Sun. A liquid xenon detector can't determine the direction of a particle that caused a signal. A detector made of diamonds could.

Walsworth envisions a detector formed of millions of millimeter-size synthetic diamonds. A diamond would flash when pierced by a neutrino or WIMP, and an automated system would remove it and scan it for an NV track, using the time of the flash to determine the track's orientation relative to the Sun and the Galaxy, the team explained last year in *Quantum Science and Technology*. Walsworth hopes to build a prototype detector in a few years. "I absolutely do not want to claim that our idea would work or that it's better than other approaches," he says. "But I think it's promising enough to go forward."

Physicists have proposed many other ideas for using quantum sensors to search for dark matter, and the influx of money should help transform them into new technologies, Zurek says. "Things can move faster when you're funded," she says. As tool builders, dark matter hunters embrace that push. "They have a great hammer, so they started looking for nails," Walsworth says. Perhaps they'll bang out a discovery of cosmic proportions. ■



A chip that could sense dark photons (top) and an axion detector, HAYSTAC, could fit on a tabletop despite their high sensitivity.

PHOTOS: (TOP TO BOTTOM) ROGER ROMANI/UNIVERSITY OF CALIFORNIA, BERKELEY; KARL VAN BIBBER

OCEANOGRAPHY

A stark future for ocean life

Model predicts a mass extinction event in the oceans if climate change is uncurbed

By **Malin L. Pinsky** *and* **Alexa Fredston**

The year 2021 marked the highest temperature and likely the lowest oxygen content for the oceans since human records began (*1*, *2*). These changes have put marine species on the front lines of climate change. For example, marine species' geographical ranges are shifting faster and experiencing more contractions than those of terrestrial species (*3*, *4*). However, whether climate change poses an existential threat to ocean life has been less clear. Marine species are often considered to be more resilient to extinction than terrestrial ones, and human-caused global extinctions of marine species have been relatively rare (*5*). On page 524 of this issue, Penn and Deutsch (*6*) present extensive modeling to reveal that runaway climate change would put ocean life on track for a mass extinction rivaling the worst in Earth's history. Furthermore, they reveal how keeping global warming below an increase of 2°C compared with preindustrial levels could largely prevent these outcomes.

The topic of climate change and species extinction on land has been fraught with controversy. This is in part because of debates over suitable methods of predicting extinctions and in part because of the relatively few documented extinctions to date (*7*). The marine research community has largely avoided making projections of extinction risk (*8*), even though experts widely see climate change as a major threat to the global oceans (*9*). This has left the watery 70% of Earth's surface as a giant blank spot in the future projection of life on Earth.

Penn and Deutsch modeled suitable habitats for marine species on the basis of well-described physiological processes that link metabolic demand for oxygen to the supply of oxygen to organismal tissues as a function of temperature. As warming causes the demand for oxygen to exceed supply in a given location, survival likely becomes untenable, causing extinctions. The authors calibrated their model against the oceanographic changes they reconstructed for the end-Permian mass extinction event, which was a period of extensive warming and deoxygenation 250 million years ago, colloquially known as the Great Dying. Although not a perfect analog to the current climate situation, the end-Permian mass extinction is one of the most cataclysmic periods in Earth's history for which there are records of extensive warming and extinction. Penn and Deutsch built confidence in their physiological model the same way climate modelers do—by integrating well-understood mechanisms and by demonstrating its ability to explain historical patterns.

The model predicts several key patterns in future extinction risk. For instance, the tropical oceans are expected to lose the most species as temperature and oxygen conditions exceed the ranges tolerable for most extant species today. However, many tropical species are expected to migrate and survive at higher latitudes as the global oceans warm up. This process mirrors the widespread range shifts that have already been observed

Department of Ecology, Evolution, and Natural Resources, Rutgers University, New Brunswick, NJ 08901, USA.
Email: malin.pinsky@rutgers.edu

PHOTO: PASCAL KOBEH/MINDEN PICTURES

Accelerating ocean warming and deoxygenation threaten a mass extinction rivalling the worst in Earth's history, especially for cold water species, such as the Atlantic rock crab shown here.

on land and in the oceans (*3*). However, polar species are likely to go globally extinct as their suitable oceanographic conditions disappear entirely. Climate change is, in effect, walking species off the ends of the Earth. Widespread warming and deoxygenation in the end-Permian mass extinction had a similar geographic signature (*10*).

The alternative scenarios explored by Penn and Deutsch illustrate the key choice that society is facing. Unabated greenhouse gas emissions would push radiative heating on Earth to 8.5 W/m^2 by 2100 and beyond in the centuries after. This future produces a doomsday scenario of end-Permian–like mass extinction by the year 2300. An alternative, consistent with the Paris Agreement (the 2016 international treaty on climate), is to limit or even reverse greenhouse gas emissions and keep radiative forcing to 2.6 W/m^2 by 2100. Doing so would keep climate change as a relatively minor threat to marine species' existence—lower than other more-direct threats, such as habitat loss or fishing—although the widespread extirpation of species from their historical locations will remain inevitable. Fortunately, greenhouse gas emissions are not on track for the worst-case scenario given policies to limit greenhouse gas emissions and the slower-than-projected growth of global economies (*11*). How close to the best-case scenario human society can hew, however, remains one of the most pressing questions for the future of life in the oceans.

The model of Penn and Deutsch also illuminates key scientific uncertainties that surround the fate of marine life, including the degree to which species can expand into new territory as it becomes climatically suitable. Observations of rapidly expanding range edges make it clear that a pessimistic, no-dispersal scenario is unrealistic, but neither is it likely that all species will perfectly track suitable habitats in an optimistic, perfect-dispersal scenario (*12*). The transient dynamics of species expansion and contraction (*13*) will likely determine which species survive in the coming decades. Another key uncertainty is the percentage of habitat that species can lose before going extinct. Calibrating against end-Permian data, Penn and Deutsch predict that loss to be 50 to 70%. Although broadly consistent with the criteria published by the International Union for Conservation of Nature (*14*), very few data exist for marine species on which to calibrate these values because most human-caused marine extinctions have been driven by overexploitation, not climate change (*15*).

The research of Penn and Deutsch also highlights key areas in need of investigation, such as metabolic traits for more marine species. Their modeling projects the future for all marine life using information from only a few dozen species, and future research will be needed to understand trait variation and correlation across life history strategies, habitats, and phylogenies. Other important extensions will include a better understanding of the role that microclimates and habitat refugia play in species survival, the potential for evolution of metabolic limits, the role of changing species interactions in promoting or impeding species survival, and the oceanographic dynamics in shelf and other coastal habitats that are currently not well represented by global climate simulations. Improved projections of these processes will refine the forecasts for marine biodiversity under climate change.

Not too long ago, canaries warned coal miners of toxic gas accumulation. Today, marine life is warning the world of a different and global gas accumulation. Staving off widespread biodiversity loss and the sixth mass extinction is a global priority. Because marine extinctions have not progressed as far as those on land, society has time to turn the tide in favor of ocean life. Exactly where the future falls between the best-case and worst-case scenarios will be determined by the choices that society makes not only about climate change, but also about habitat destruction, overfishing, and coastal pollution. With a coordinated approach that tackles multiple threats, ocean life as we know it has the best chance of surviving this century and far beyond. ■

REFERENCES AND NOTES

1. L. Cheng *et al.*, *Adv. Atmos. Sci.* **39**, 373 (2022).
2. D. Breitburg *et al.*, *Science* **359**, eaam7240 (2018).
3. J. Lenoir *et al.*, *Nat. Ecol. Evol.* **4**, 1044 (2020).
4. M. L. Pinsky, A. M. Eikeset, D. J. McCauley, J. L. Payne, J. M. Sunday, *Nature* **569**, 108 (2019).
5. D. J. McCauley *et al.*, *Science* **347**, 1255641 (2015).
6. J. L. Penn, C. Deutsch, *Science* **376**, 524 (2022).
7. O. T. Lewis, *Phil. Trans. R. Soc. B* **361**, 163 (2006).
8. M. C. Urban, *Science* **348**, 571 (2015).
9. S. Knapp *et al.*, *Sci. Total Environ.* **574**, 191 (2017).
10. J. L. Penn, C. Deutsch, J. L. Payne, E. A. Sperling, *Science* **362**, eaat1327 (2018).
11. M. G. Burgess, J. Ritchie, J. Shapland, R. Pielke Jr., *Environ. Res. Lett.* **16**, 014016 (2021).
12. A. Fredston *et al.*, *Glob. Change Biol.* **27**, 3145 (2021).
13. J. Pagel, F. M. Schurr, *Glob. Ecol. Biogeogr.* **21**, 293 (2012).
14. International Union for Conservation of Nature (IUCN), "IUCN Red List Categories and Criteria: Version 3.1" (IUCN, ed. 2, 2012).
15. N. K. Dulvy, Y. Sadovy, J. D. Reynolds, *Fish Fish.* **4**, 25 (2003).

ACKNOWLEDGMENTS

The authors acknowledge funding from the US National Science Foundation (nos. DEB-1616821 and OIA-1936950) and the Lenfest Ocean Program (no. 00032755). M.L.P. serves as an Oceana science advisor.

10.1126/science.abo4259

INFECTIOUS DISEASES

From outbreaks to endemic disease

Relief from the effects of epidemics may signal the start of low-level disease persistence

By **Michael F. Antolin**

In epidemiology, endemism describes a state in which pathogens persist at low levels even if they cause high mortality and morbidity during outbreaks. Endemism occurs when disease falls to a low and possibly tolerable (treatable) level. But persistence at low prevalence presents a health challenge, especially for virulent pathogens that may spill over to humans and companion animals, spill back to other wildlife species, or reemerge again in epidemics. On page 512 of this issue, Mancy *et al.* (*1*) present a longitudinal study (from 2002 to 2016) of canine rabies in the Serengeti region of northern Tanzania. This study addressed how the rabies virus continues to circulate in domestic dogs at less than 0.15% prevalence while showing (weak) patterns of density dependence (more rabies at higher population densities of dogs). This study highlights the continued health risks posed by pathogens that become endemic.

The dataset reported by Mancy *et al.* includes observations of tens of thousands of domestic dogs; contact tracing of infections from rabid dogs to other dogs and humans; and detailed tracking of dog densities, disease incidence, and vaccination campaigns. By implementing a complex individual-based model using the spatially detailed data, the authors determine that endemic persistence of *Rabies lyssavirus* depends on identifying the spatial scale (1 km^2) at which relatively small outbreaks begin by arrival of rabid dogs from elsewhere. A critical feature is that transmission occurs through superspreading, in which a few rabid dogs travel relatively long distances and are responsible for most of the infectious bites. Rabies outbreaks in Tanzania resolved into 22 dominant transmission chains that arose in differ-

Department of Biology, Colorado State University, Fort Collins, CO, USA. Email: michael.antolin@colostate.edu

ent localities at different times. A previous study from the same area sequenced *R. lyssavirus* isolates from a subset of cases and placed them in a phylogeographic framework that verified the existence of genetically identifiable lineages (*2*). Similar rabies transmission chains have been identified in the Central African Republic (*3*).

Canine rabies provides a clear example of a virulent pathogen, which is inevitably fatal because dogs do not naturally develop immunity, endemically persisting because transmission and infection are spatially segregated into a metapopulation (*4*, *5*). In a rabies metapopulation, local flare-ups in places with denser dog populations burn themselves out, but rabid dogs disperse while transmitting the virus and starting new local outbreaks. With rabies prevalence being low across the region of northern Tanzania, rabies infections cannot at that scale be predicted from dog densities. Disaggregating the region into individual lineages at the scale of 1 km^2 reveals locally subdivided and temporally separated infection dynamics. Furthermore, dynamics are density-dependent at the local scale, with rabies persisting in the region asynchronously and endemically as a metapopulation with multiple independent chains of transmission.

Canine rabies was tracked in tens of thousands of domestic dogs in northern Tanzania to understand the persistence of endemic disease.

Localized contagion and outbreaks of disease are not new to the human experience. This was observed in Medieval Europe during the Black Death (plague) era that began in 1347 AD and caused local outbreaks for three centuries. The advice "Fuge cito, vade longe, redetarde" (flee quickly, go far, return slowly) was commonly given when outbreaks flared up. Plague (*Yersinia pestis*) is a zoonotic, flea-borne bacterial pathogen of ground-dwelling rodents that catastrophically spilled over to humans in at least three large-scale epidemics dating back to 550 AD. Plague does not persist as an epidemic human disease and now recurs only in places where competent rodent hosts and their fleas are uncontrolled in human proximity (*6*). Persistence in rodent host populations, however, depends on metapopulation dynamics (*7*–*9*). For example, plague outbreaks among prairie dog colonies on the Great Plains of the United States sporadically cause local extinctions of prairie dog towns, which are then recolonized ~4 to 5 years after the epidemic wanes (*8*). Plague retains high virulence and, in the time since its introduction into North America in the early 1900s, has drastically reduced overall prairie dog abundance. Despite high local mortality, plague has not caused overall extinction of prairie dogs from the grasslands. Like rabies, plague persists endemically and is usually undetectable between outbreaks in prairie dog towns.

Does rabies endemism in dogs in Tanzania represent a general phenomenon of pathogen persistence? Rabies in Tanzania is in many ways a distinct example, with transmission directly from dog to dog, no wild animal reservoir, a long serial interval between infections (up to several hundred days), and infections that are nonimmunizing and inevitably fatal. Other pathogens have a diversity of transmission modes (directly among hosts, vector-borne, or from animal or environmental reservoirs), differences in serial infection intervals (time from transmission to a host until the onset of symptoms), and heterogeneity of transmission patterns (superspreaders) and can elicit variable immune responses. All of these factors can affect persistence and endemism. Further, in host-pathogen systems in the wild, observing infection dynamics through contact tracing, accurately estimating host abundance, evaluating host immunity, and detecting movements of infected hosts without bias is much more difficult (*10*).

What is generalizable is that pathogen transmission is scale-dependent and that most host populations are subdivided and are not well mixed at large scales. Ecological interactions, including pathogen transmission patterns, generally occur locally (*10*–*12*), and epidemiological studies that aggregate data to large spatial scales can result in spurious conclusions (*13*). Mancy *et al.* tested alternative spatial scales in their model, with the local scale of 1 km^2 providing the best fit to the rabies prevalence data.

After 2 years of the COVID-19 pandemic, a transition to endemism and the end of widespread infection chains and less obtrusive public health measures may be in reach. Widespread vaccination should lessen disease and slow emergence of new viral variants. The likelihood, however, is that severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) will not entirely burn out but will slide into endemic persistence in localized chains of transmission. But endemic persistence still carries health risks in local populations. In Tanzania during the 2002 to 2016 period, more than 1000 humans were bitten by rabid dogs, with 44 rabies deaths (*1*). Vaccination of dogs has been implemented but not at high enough levels to stop transmission and eradicate *R. lyssavirus*. Even in the modern vaccination era, several early-childhood diseases (measles, mumps, rubella, and whooping cough) and the dengue virus persist in local transmission chains that escape detection without regular surveillance (*14*, *15*). Although endemism is more desirable than an epidemic, it is not the end of disease but a new and alternative challenge. ■

PHOTO: KATE HAMPSON

REFERENCES AND NOTES

1. R. Mancy *et al.*, *Science* **376**, 512 (2022).
2. K. Brunker *et al.*, *Mol. Ecol.* **27**, 773 (2018).
3. H. Bourhy *et al.*, *PLOS Pathog.* **12**, e1005525 (2016).
4. G. Hess, *Ecology* **77**, 1617 (1996).
5. B. Grenfell, J. Harwood, *Trends Ecol. Evol.* **12**, 395 (1997).
6. K. L. Gage, M. Y. Kosoy, *Annu. Rev. Entomol.* **50**, 505 (2005).
7. M. J. Keeling, C. A. Gilligan, *Nature* **407**, 903 (2000).
8. S. Davis *et al.*, *Science* **304**, 736 (2004).
9. D. J. Salkeld *et al.*, *Bioscience* **66**, 118 (2016).
10. J. Antonovics, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **372**, 20160087 (2017).
11. J. F. Addicott *et al.*, *Oikos* **49**, 340 (1987).
12. J. N. Thompson, *The Geographic Mosaic of Coevolution* (Univ. Chicago Press, 2005).
13. D. J. Salkeld, M. F. Antolin, *EcoHealth* **17**, 4 (2020).
14. C. J. E. Metcalf, K. Hampson, A. J. Tatem, B. T. Grenfell, O. N. Bjørnstad, *PLOS ONE* **8**, e74696 (2013).
15. H. Salje *et al.*, *Science* **355**, 1302 (2017).

10.1126/science.abo7428

BATTERIES

A granular approach to electrode design

The consistency of cathode particles plays a pivotal role in battery performance

By **Jie Xiao**

Tailoring the intrinsic properties of battery materials and designing their multiscale structures are very important to maximize their electrochemical performance. When electrochemically active materials—often in powder form—are used in batteries, they are usually dispersed within porous composite electrodes immersed in a liquid electrolyte and provide ionic and electronic conduction pathways for the electrochemical reactions. Therefore, the full utilization of electrochemically active particles in the electrodes (i.e., cathodes and anodes) depends not only on their own properties but also on those of the conductive network for the particles, as well as the interactions between the particles and the network. On page 517 of this issue, Li *et al.* (*1*) report on the dynamics of particle networks in composite cathode for lithium-ion batteries and highlight the importance of size consistency for the electrode particles.

The functionality of an electrode depends on the integration of the respective contributions from every active particle in the electrode network. A typical battery electrode consists of electrochemically active particles embedded in a porous composite whose porosity can be tailored for specific applications (*2*). Ideally, the electrolyte infiltrates into every pore of the electrode to ensure good ionic conductivity. In the perfect scenario, ions are accessible everywhere, and the conductivity within the electrodes is good because of the interconnected carbon in the electrode structure (*3*). However, in reality, the local conductive network surrounding the active cathode particles, which facilitates the interactions of ions and electrons for the redox reactions, is not homogeneous, and thus the reaction rate of the active particles is also not homogeneous.

The electrode heterogeneity creates a dynamic evolution of the electrochemical activities of individual particles inside the electrode, which affects battery performance. The heterogeneous and dynamic electrochemical activities of electrode particles cannot be measured by the electrochemical output, such as the voltage or the battery capacity, because these bulk measurements do not give information about the movement and interactions among the individual particles. During charge-discharge cycles, when lithium ions go in and out of the host particles in the cathode, a dynamic equilibrium of the particle system is reached through the coevolution of all particles driven by spontaneously yielded internal pressure. Tracking the interactions of individual active particles within the electrode is needed to understand the dynamic equilibrium within the electrode.

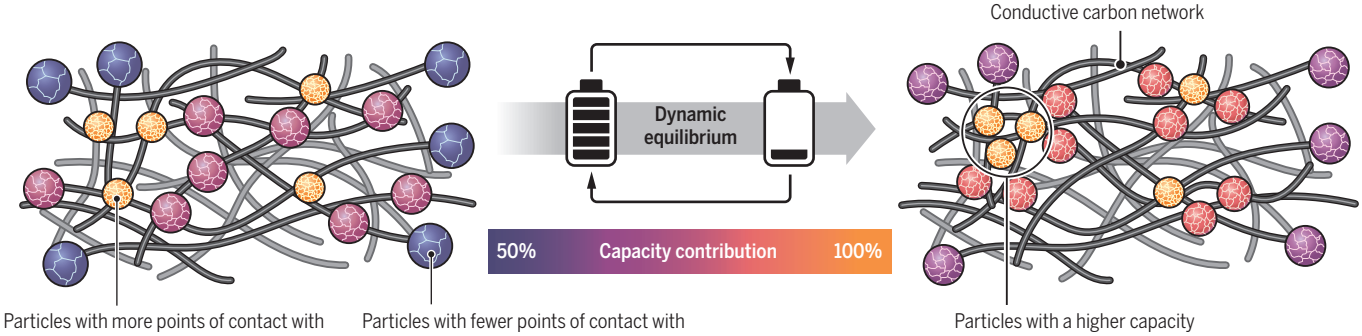
Li *et al.* investigated the dynamic evolution of the particles in a battery cathode (see the figure). Using synchrotron nanoresolution holo-tomography, they tracked the morphological and chemical characteristics of thousands of individual active particles. Inside the cathode, the morphological defects and the electrochemical activities of individual particles coevolved and transitioned through three distinct stages: single-particle activation, where electrode particles begin participating in the electrochemical reactions; followed by segregation of groups of particles, where the utilization of particles and degree of damage to the particles vary; and finally, global homogenization, where the modulation between the electrochemical activity and mechanical damage reduces those variations. The authors also observed a correlation between the collective electrochemical activity of, and the damage to, the particles that occur during charge-discharge cycles. According to their model, particles with higher electrochemical activity are more likely to experience mechanical damage, such as particle fracture or detachment from the electrodes. This mechanical damage in return would suppress the electrochemical reaction by increasing the resistance.

The findings of Li *et al.* highlight the need for high-throughput analysis and the importance of morphological and chemical consistencies of electrochemically active particles for electrode robustness against charge-discharge cycles. The relevance of

Pacific Northwest National Laboratory, Richland, WA 99352, USA. Email: jie.xiao@pnnl.gov

The evolution of electrochemically active particles in composite cathodes

Each active cathode particle evolves and shifts inside composite cathodes differently during charge-discharge cycles. This is caused by the heterogeneity of the conductive structure (shown here as carbon) and various inconsistencies among the particles, such as in size and degree of cracking. Over time, the particles settle into clusters and stabilize the electrochemical property of the battery.


Conductive carbon network
Dynamic equilibrium
50%
Capacity contribution
100%
Particles with more points of contact with the conductive carbon have more cracking and a higher capacity contribution.
Particles with fewer points of contact with the conductive carbon have less cracking and a lower capacity contribution.
Particles with a higher capacity contribution clump together, affecting the global homogenization transition.

GRAPHIC: KELLIE HOLOSKI/*SCIENCE*

particle homogeneity for battery materials has been well realized in commercial batteries. In some cases, a mixture of particles with different sizes is purposely packed into electrodes to increase the electrode density, provided that the particles are smaller than the critical size beyond which cracking would occur (*4*, *5*). In academic research, however, there are often considerable variations in the size, shape, and surface properties of the same materials being used by industry, which can lead to inconsistent findings. Li *et al*. highlight the value of selecting materials with particle-level consistency for application-oriented research.

The contact between active particles and carbon, as well as the material consistency of the porous electrode itself, affect the capacity and cycle life of batteries. Research efforts using field-guided self-assembly (*6*), freeze-drying (*7*), or printing technologies (*8*) to gauge the electrode-level structure ordering are possible solutions if the porosity of electrodes can be better controlled and monitored. The electrochemical performance of a battery is determined by the electrode that's the weakest link. When assessing cathode (or anode) materials, the corresponding anode (or cathode) must also be sufficiently stable that the observed battery performance is not muddled by unwanted discrepancies.

In the future, it would be of great interest to develop an integrated approach for studying all the different components of a battery. For example, a carbon additive in positive electrodes is usually only about 2% by weight but establishes the entire conductive carbon framework for electrodes (*3*). Similarly, the distribution of the very limited amount of electrolyte in a battery plays an equally outsized role in battery performance, specifically, in how ions move within the electrodes. A fundamental understanding and large-set data analysis of the dynamics of the particulate system in battery electrodes can help guide more refined battery materials and electrode manufacturing in the future, which will only become more relevant as the world gradually moves away from fossil fuel dependency. ■

REFERENCES AND NOTES

1. J. Li *et al*., *Science* **376**, 517 (2022).
2. J. Xiao *et al*., *J. Electrochem. Soc.* **169**, 010524 (2022).
3. Y. Bi *et al*., *J. Electrochem. Soc.* **169**, 020521 (2022).
4. Y. Bi *et al*., *Science* **370**, 1313 (2020).
5. H. Kim *et al*., *Angew. Chem. Int. Ed.* **49**, 2146 (2010).
6. D. Wang *et al*., *ACS Nano* **4**, 1587 (2010).
7. H. Shen *et al*., *ACS Appl. Mater. Interfaces* **12**, 3494 (2020).
8. T. S. Wei *et al*., *Adv. Mater.* **30**, 1703027 (2018).

10.1126/science.abo7670

Honey bees initiating dance language experience an increased concentration of dopamine, indicating the presence of a wanting system in insects that is analogous to that observed in mammals, including primates.

PHOTO: KIM TAYLOR/MINDEN PICTURES

NEUROSCIENCE

Why do animals want what they like?

As in mammals, honey bee motivation for wanting rewards is modulated by dopamine

By **Jair E. Garcia**[1] *and* **Adrian G. Dyer**[1,2]

The ability to experience pleasure or displeasure is created by the brain and therefore has a physiological component (*1*). In 1872, based on his observations of humans and animals, Charles Darwin proposed that evolution may select affective reactions that shape behavior, and thus individual fitness (*2*). Affective neuroscience in mammals searches for mechanistic explanations that underpin the experience of liking or disgust (*1*). Recently, the role of dopamine was reevaluated, showing that it is an important neuromodulator for wanting rather than liking rewards (*3*). On page 508 of this issue, Huang *et al*. (*4*) show that regulation of the neurotransmitter dopamine is also an important component of the motivation of wanting in honey bees (*Apis mellifera*). This suggests that the fitness benefits of a motivation wanting system regulated by dopamine are likely to be conserved and may explain behavior in a wide range of animals.

In humans and other mammals, analogous brain circuitry is activated by the wanting or desire for a reward (*1*). For example, different reward types such as food, sex, addictive drugs, and even various forms of experienced artwork show evidence of stimulating a brain network that includes the prefrontal cortex, ventral pallidum, and amygdala of the primate brain in a similar manner. Additionally, common mammalian brain structures, such as the amygdala in rodents (*3*), also show activation to wanting a reward (*5*). This suggests a potential common reward network of interacting brain regions that underpin rewarding experiences. An important component of such a reward system is the regulation of dopamine, which helps modulate incentive motivation, the wanting for certain rewards.

Honey bees are an important comparative species for how brains learn, and together with other insect models, such as the common fruit fly (*Drosophila melanogaster*), are a test bed to investigate what might be the minimum brain structures required for infor-

[1]BIDs Lab, School of Media and Communication, RMIT University, Melbourne, VIC, Australia. [2]Department of Physiology, Monash University, Clayton, VIC, Australia. Email: adrian.dyer@rmit.edu.au; adrian.dyer@monash.edu

mation processing (*6*). Honey bees are adept learners of a variety of visual and olfactory stimuli when associated with a sweet-tasting sugary reward. Individual bees are then motivated to communicate the value of their perceived experience to hive mates through the bee waggle dance language (*7*). This behavioral paradigm enabled decoding of the bee dance language that contains information about direction and distance of profitable food sources and enables good experimental access for testing bee sensory perception (*8*).

Huang *et al.* investigated how honey bees forage, and what might motivate individuals to leave the hive in search of sweet-tasting rewards in the environment. By monitoring the foraging and subsequent dance language by bees from different test groups that were captured at different phases of their activities, it was possible to understand the potential role of the dopaminergic system in the anticipation of wanting a reward. When bees that had recently visited a profitable food source started dancing to communicate to other bees in the hive, the dancing bees showed a significant increase in dopamine brain concentrations, which was quantified using high-performance liquid chromatography on single brains. However, this increase is temporary and drops by the end of the bee's waggle dance. When dancing bees subsequently depart the hive to seek more rewards, dopamine again increases, followed by a decrease while feeding. These fast transient changes of dopamine concentrations in the bee brain show that it is the expectation of wanting a reward that drives this important component of how motivation is modulated in the reward system—a mechanism similar to that reported in some mammals (*1*, *5*).

The role of dopamine as a biologically relevant component of foraging motivation was confirmed in odor-learning experiments in which bees were topically treated with increasing dopamine amounts and subsequently showed significantly better learning profiles compared to controls. Bees treated with exogenous dopamine also demonstrated improved memory for recalling learned tasks compared to controls. Thus, modulation of dopamine concentrations in bee brains is essential in a variety of foraging tasks, including expectation, motivating individuals, and improving their capacity to learn important flower signaling such as odor to identify profitable foraging options (*9*). An improved understanding of the mechanisms that mediate expectation, wanting, and motivation in bees (and likely many other animals) according to individual experience and resulting dopamine concentrations in the brain could also explain why behavioral differences between individuals are frequently reported, even for individuals from the same hive (*10*).

Results from the study of Huang *et al.* indicate that honey bees share with mammals common neural mechanisms for encoding the wanting of stimuli with positive hedonic value. Thus, as anticipated by Darwin 150 years ago, evolution likely shaped wanting mechanisms that improve animal fitness, by experiencing pleasure from nutritious food, or triggering responses from emotional stimuli linked to survival brain circuits (*11*). Conversely, an imbalance of this mechanism may have negative effects on fitness by, for example, promoting pathologies such as addiction (*1*).

Honey bees and a variety of native bees are major pollinators of important crops, including almonds, citrus fruits, and many vegetables, with both human nutritional and economic value in the tens of billions of dollars a year within the US alone (*12*). A deeper understanding of how dopamine modulates wanting in these important animals can lead to improved ways of motivating bees to forage. For example, the queen bee in a hive is able to deploy chemical pheromones that will capture the attention of young bees to entice them to perform certain tasks by stimulating dopamine pathways in the brain (*13*), and equivalent stimulations to drive the wanting systems of forager bees would likely benefit pollination services by motivating them to visit more flowers.

As the honey bee continues to be a key model for what can be achieved with a miniaturized brain (*6*), understanding the neural mechanisms of wanting and the evidence of variability among individuals (*10*), depending upon acquired experience, is informative for explaining different animal behaviors. ■

REFERENCES AND NOTES

1. K. C. Berridge, M. L. Kringelbach, *Curr. Opin. Neurobiol.* **23**, 294 (2013).
2. C. Darwin, *The Expressions of Emotions in Man and Animals* (Oxford Univ. Press, 1872).
3. K. C. Berridge, M. L. Kringelbach, *Neuron* **86**, 646 (2015).
4. J. Huang *et al.*, *Science* **376**, 508 (2022).
5. K. C. Berridge, T. E. Robinson, *Am. Psychol.* **71**, 670 (2016).
6. R. Menzel, *Nat. Rev. Neurosci.* **13**, 758 (2012).
7. K. von Frisch, *The Dance Language and Orientation of Bees* (Harvard Univ. Press, 1967).
8. M. V. Srinivasan, S. Zhang, M. Altwein, J. Tautz, *Science* **287**, 851 (2000).
9. A. Kantsa *et al.*, *Nat. Ecol. Evol.* **1**, 1502 (2017).
10. L. Chittka, P. Skorupski, N. E. Raine, *Trends Ecol. Evol.* **24**, 400 (2009).
11. J. LeDoux, *Neuron* **73**, 653 (2012).
12. S. G. Potts *et al.*, Summary for policymakers of the assessment report of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services on pollinators, pollination and food production (IPBES, 2016).
13. K. T. Beggs *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 2460 (2007).

ACKNOWLEDGMENTS

A.G.D. receives funding from the Australian Research Council.

10.1126/science.abp8609

MATERIALS SCIENCE

An adsorbent with flexible nanoscopic pores

Scanning transmission electron microscopy shows the adaptive pores of a zeolite

By **Tom Willhammar** *and* **Xiaodong Zou**

Nanoporous materials have large surface areas and well-defined pores at the molecular scale, making them attractive as selective adsorbents and catalysts (*1*). They can act as molecular sieves and have garnered interest because of their potential use as energy-efficient adsorbents for gas separation and storage applications, such as carbon capture. The adsorption properties of some of these materials suggest that interactions between the nanoporous material and the molecules or ions it adsorbs might induce a degree of structural flexibility in the pores. On page 491 of this issue, Xiong *et al.* (*2*) describe changes in the pore geometry of the industrially important zeolite ZSM-5 during the gas adsorption-desorption process. This direct observation of structural flexibility in a nanoporous material reveals ways to manipulate the dynamic behavior and function of these materials.

Zeolites have a nanoporous three-dimensional structure built from corner-sharing silica (SiO_4) and alumina (AlO_4) tetrahedra. To date, more than 250 different zeolite structures have been discovered (*3*), each with a distinctive pore size and shape. However, adsorption experiments in these materials have shown that gas molecules larger than the expected pore size can enter the material, which indicates structural flexibility of the material associated with the so-called "host-guest" interactions. To understand the size and shape selectivity properties of a material as an adsorbent

Department of Materials and Environmental Chemistry, Stockholm University, SE-106 91 Stockholm, Sweden. Email: tom.willhammar@mmk.su.se; xzou@mmk.su.se

or catalyst, it is important to know how its structure changes during the adsorption-desorption of the gas molecules.

Structural flexibility caused by physical or chemical stimuli has been reported in some zeolites (*4*). This structural flexibility can be induced by ion exchange, dehydration, or gas adsorption, and the flexibility can manifest as a change in the unit-cell volume and/or in the symmetry and pore size. Diffraction-based techniques, including x-ray, neutron, and electron diffraction, have served as important tools for studying these structural flexibilities. For example, in situ powder x-ray diffraction has shown that the unit-cell volume can expand by up to 25% upon ion exchange (*5*). The accessible pore window can be substantially altered upon dehydration or ion exchange, which greatly affect the adsorption and desorption properties (*4*, *6*). Three-dimensional electron diffraction (3DED) has also been a powerful tool for determining the structure of polycrystalline porous materials (*7*). Researchers have used 3DED to locate carbon dioxide (CO_2) molecules in the zeolite CHA (*8*). However, these diffraction-based techniques can only provide structure information averaged over the entire measured material. To obtain a more detailed understanding of the mechanisms underlying structural flexibility, it is necessary to visualize the host-guest interactions at the individual-pore level.

Recent advances in imaging techniques with atomic resolution have made it possible to visualize the guest molecules inside the voids of porous materials. One such method is integrated differential phase contrast scanning transmission electron microscopy (iDPC-STEM). Researchers have used iDPC-STEM to observe physically adsorbed organic molecules and chemically bound single molybdenum oxide clusters in the nanopores of zeolite ZSM-5 (*9*). When synthesizing zeolites, organic molecules are often used as a "structure-directing agent" to help build the three-dimensional structure of a zeolite. Using iDPC-STEM, researchers have been able to track the location of the structure-directing organic molecules in the zeolite pores, which provides insights into the formation of the zeolite material (*10*). iDPC-STEM imaging has also been used to observe para-xylene molecules adsorbed inside the straight channels of the zeolite MFI (*11*). This revealed that the orientation of the para-xylene molecules in the channel is correlated to the channel geometry and therefore could be used to detect host-guest van der Waals interactions in porous materials (*11*).

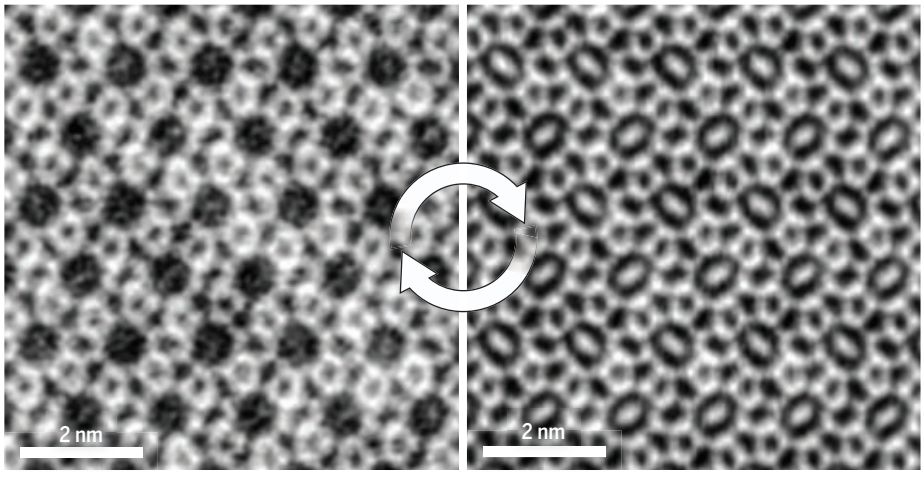
Using an in situ transmission electron microscopy gas-cell specimen holder, Xiong *et al.* studied the gas adsorption-desorption process in zeolite ZSM-5 using benzene as a probe molecule (see the figure). Their imaging results show a reversible local deformation of pore windows from circular to elliptical upon the pressure-induced adsorption-desorption of the benzene molecules. The pores are stretched by up to 15% in aspect ratio to optimize the interaction with the benzene molecules and facilitate their diffusion into the channels. The pore structural flexibility is associated with the flexibility of the Si–O–Si bond angles, and the introduction of benzene molecules in the pores reduces such flexibility, as predicted by ab initio molecular dynamic simulation. The direct observation of such a delicate interplay between the zeolite host and guest molecules allows a better understanding of the mechanisms behind the complex adsorption-desorption behavior, which will help in the design and selection of adsorbents for specific applications.

Unlike zeolites, metal-organic frameworks (MOFs) have a higher degree of structural flexibility. Structural changes in MOFs can be triggered using different guest molecules for different adsorption-desorption behaviors, such as reversible pore opening-shrinking, or "breathing," and selective gate-opening (*12*), which has been studied using x-ray techniques. In earlier efforts to study the CO_2 adsorption of MOFs, researchers have used low-dose cryogenic electron microscopy to observe CO_2 molecules in the channels of the MOF ZIF-8 (*13*). Direct observation of sorption-induced structural changes of the MOFs at the individual-pore level could be achieved by a combination of an in situ gas environment and low–electron dose imaging to prevent electron beam damage to the MOFs. This knowledge is important for the design of MOFs where structural flexibility can be used to tailor adsorption properties through chemical or physical interactions between the host and guest.

The possibility of visualizing local structural changes induced by adsorbed molecules during an entire adsorption-desorption process should serve as an inspiration for applications of in situ electron microscopy on other flexible nanoporous materials beyond zeolites and MOFs. Direct observation of guest molecules in the channels using iDPC-STEM imaging will provide an understanding of host-guest interactions in all those nanoporous materials, which is important for developing efficient adsorbents and catalysts with enhanced selectivity and performance. ■

The flexible pores of a zeolite

The nanopores of a zeolite adapt their shape to fit the adsorbed molecules. Using scanning transmission electron microscopy, Xiong *et al*. directly observed this structural change at the sub–unit cell scale during the entire adsorption-desorption process.



GRAPHIC: BASED ON (2)

REFERENCES AND NOTES

1. M. E. Davis, *Nature* **417**, 813 (2002).
2. H. Xiong *et al.*, *Science* **376**, 491 (2022).
3. C. Baerlocher, L. B. McCusker, Database of zeolite structures; http://www.iza-structure.org/databases/.
4. M. M. Lozinska *et al.*, *J. Am. Chem. Soc.* **134**, 17628 (2012).
5. D. R. Corbin *et al.*, *J. Am. Chem. Soc.* **112**, 4821 (1990).
6. J. Grand *et al.*, *Chem. Mater.* **32**, 5985 (2020).
7. Z. Huang et al., *Chem. Sci.* **12**, 1206 (2021).
8. M. Debost *et al.*, *Angew. Chem. Int. Ed.* **59**, 23491 (2020).
9. L. Liu *et al.*, *Angew. Chem. Int. Ed.* **59**, 819 (2020).
10. J. Cho *et al.*, *Chem. Mater.* **33**, 4146 (2021).
11. B. Shen *et al.*, *Nature* **592**, 541 (2021).
12. A. Schneemann *et al.*, *Chem. Soc. Rev.* **43**, 6062 (2014).
13. Y. Li *et al.*, *Matter* **1**, 428 (2019).

10.1126/science.abo5434

INFECTIOUS DISEASE

Resurgence of avian influenza virus

Unprecedented outbreaks of the H5N1 highly pathogenic avian influenza virus raise concern

By **Michelle Wille**[1,2,3] and **Ian G. Barr**[2,3]

Highly pathogenic avian influenza viruses (HPAIv) have recently been detected on the East coast of Canada (December 2021) and the United States (January 2022) and have been found in both wild birds and poultry (*1*). This comes after the same strain (known as H5N1) swept through Asia, Africa, and Europe in late 2021, replacing the previous HPAIv and causing widespread outbreaks and millions of deaths in poultry and wild birds. These HPAIv are of concern not only to birds but also to humans because they pose a potential pandemic risk. How has this virus emerged and spread so rapidly, and what does it mean for poultry, wild birds, and humans?

Influenza A viruses infect humans and a wide range of animal species such as pigs, poultry, and many wild birds, including waterbirds, the natural hosts of these viruses. Influenza A viruses are classified based on differences in their two most abundant surface proteins, hemagglutinin (H) and neuraminidase (N), of which there are 16 and 9 different subtypes, respectively, that infect avian species; these occur in different combinations, defining the virus (e.g., H5N1). Wild birds are infected with low-pathogenic avian influenza viruses (LPAIv), which they can carry asymptomatically. Only two H subtypes (H5 and H7) have repeatedly emerged as HPAIv to date, and these usually occur when LPAIv H5 and H7 viruses cross from wild birds into poultry, where changes in the H protein transform them into HPAIv, causing severe disease, devastating outbreaks, and up to 100% mortality in chickens (*2*). Although the N subtype may play some role in virus transmissibility, it does not determine disease severity.

The current circulating lineage of H5 HPAIv had its beginnings in Asia as early as 1996 as the goose/Guangdong lineage (gs/GD). After the emergence of a novel lineage and associated outbreaks in 2003, it spread to Asia, the Middle East, Africa, and Europe, where >400 million chickens, turkeys, and domestic ducks were culled and >600 human cases of H5N1 HPAIv were recorded within 5 years (*3*). These viruses became endemic in the poultry of many countries in Asia and Africa, with the gs/GD lineage diversifying into several genetic clades as a result of high infection rates, vaccination pressure, and circulation of this virus in an array of avian hosts. Since 2014, after evolution of the H gene, a new gs/GD lineage designated 2.3.4.4 has emerged, and these HPAIv have subsequently dominated outbreaks globally.

Why has the H5N1 subtype combination become globally dominant again? The reasons for the current 2021–2022 wave of H5 HPAIv are likely to be multifactorial. The reversion back to using N1 is probably serendipitous, but the current H5N1 virus clearly has a fitness advantage over the previous H5N6 or H5N8 HPAIv that predominated from 2014 to 2021. This has allowed H5N1 to rapidly displace other H5Nx viruses. Normally such a dramatic global sweep would be caused by a major change in the influenza H protein, but the H of the current H5N1 is similar to that of the previous H5N6 and H5N8 viruses, falling into the same genetic lineage (2.3.4.4). This leaves properties in the N1 or changes to other viral genes through reassortment or mutation as possible reasons for increased fitness. However, there may be other host factors leading to this spread, such as being able to infect a broader range of wild birds or reaching higher viral loads in birds, causing larger and more intense outbreaks. These matters need further, careful investigation.

How has the rapid spread of H5N1 HPAIv occurred? Although HPAIv often emerge in high-density poultry production systems, it is wild birds that can spread these viruses globally. The role of wild birds in spreading HPAIv H5 before 2014 was unclear. However, in late 2014, wild waterfowl spread HPAIv H5Nx from their breeding areas in Siberia to Europe and North America (*4*) and to Africa in 2017 along their established migratory flyways. Wild birds have now reintroduced H5 HPAIv into North America with detections in wild birds: great black-backed gulls (*Larus marinus*) in St. John's Canada (November to December 2021) (*4*) and wild birds as well as backyard and commercial poultry flocks in the United States from January 2022 (*1*) (see the figure). H5Nx HPAIv have now been introduced into all continents except Australia, South America, and Antarctica, which have probably been spared because of ecological barriers to waterfowl migration from areas of endemic circulation; however, this situation could change.

What are the consequences of the H5N1 HPAIv spread? Although wild birds are able to spread these viruses through migration, the viruses may also cause their own mass mortality events. These mortality events in wild birds come with substantial conservation concern. For example, H5N1 HPAIv caused the death of hundreds of red knots (*Calidris canutus*) in the Netherlands in 2021 (*5*), 10,000 migratory common cranes (*Grus grus*) in Israel in December 2021 (*6*), ~10% of the Svalbard breeding population of barnacle geese (*Branta leucopsis*) in late 2021 (*7*), and hundreds of Dalmatian pelicans (*Pelecanus crispus*) in northern Greece in 2022.

Outbreaks of H5Nx HPAIv have also had a substantial impact on poultry production and cost the sector billions of dollars since their emergence. Data reported to the World Organisation for Animal Health (OIE) found that between 2005 and 2019, there were 18,620 outbreaks in poultry reported across 76 countries. In 2020–2021 alone there have been more than 3000 HPAIv events reported (*8*), with ~15,000,000 poultry losses (died or culled) globally as a result of H5Nx. The United Kingdom has experienced its biggest outbreak due to H5N1 HPAIv, with millions of poultry culled; there have also been 2804 outbreaks in Europe since October 2021 (up to 23 March 2022) (*8*, *9*).

Onward transmission of HPAIv to humans is an ongoing threat that continues to be of pandemic concern. From 2003 to 2021, there have been 863 human cases of H5N1 HPAIv, resulting in 456 deaths reported to the World Health Organization (WHO), with virtually all of these cases occurring before 2016. Although this earlier "classic" strain of H5N1 caused severe infections in more than half of those infected, the number of people infected was relatively low given that millions of birds and hundreds of thousands of poultry workers, farmers, and live-bird market-stall holders were in direct contact with infected birds. Furthermore, these infections were almost exclusively associated with contact with infected birds, with no sus-

[1]Sydney Institute for Infectious Diseases, School of Life and Environmental Sciences and School of Medical Sciences, The University of Sydney, Sydney, New South Wales, Australia. [2]World Health Organization (WHO) Collaborating Centre for Reference and Research on Influenza, The Peter Doherty Institute for Infection and Immunity, Melbourne, Victoria, Australia. [3]Department of Microbiology and Immunology, The Peter Doherty Institute for Infection and Immunity, The University of Melbourne, Melbourne, Victoria, Australia. Email: michelle.wille@sydney.edu.au; ian.barr@influenzacentre.org

tained onward human-to-human transmission detected, confirming that these viruses were not well-enough adapted to humans to be a major concern. The more recent 2.3.4.4 H5 HPAIv have produced similarly low rates of human infections, possibly because these viruses share the same ancestral gs/GD lineage. However, different pathogenicity has been observed when human H5 infections do occur, with the earlier H5N1 and H5N6 HPAIv appearing more severe and causing more deaths in humans than H5N8 viruses. Notably, all these H5Nx HPAIv strains remain highly transmissible and deadly to poultry and, to a lesser extent, wild birds.

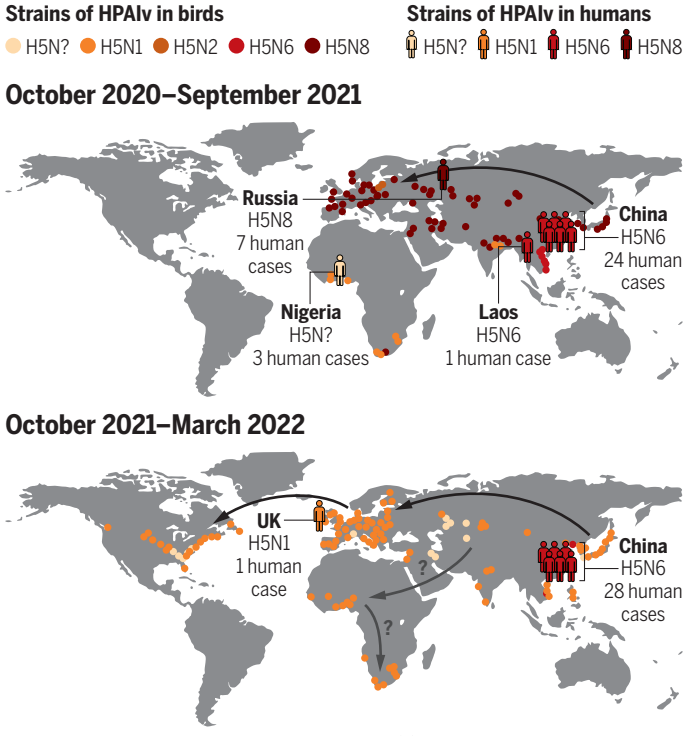
H5N8 2.3.4.4 HPAIv have resulted in very few human cases despite their high circulation from 2014 to 2021. Several mild H5N8 infections were detected in December 2020 in poultry workers in Astrakhan Oblast (Russian Federation), none of whom required hospitalization (*10*). Three human infections with H5N? HPAIv (probably N1) were reported in Nigeria in March 2021, but it is unclear if these were true infections or environmental carriage of virus because all cases were asymptomatic. By contrast, between 2014 and 16 March 2022, H5N6 has caused 77 human infections with 32 deaths, with all but one case occurring in China. Of recent concern has been the apparent increase in H5N6 HPAIv human cases in China in 2021, where 33 cases (out of 76 total cases) with 11 deaths were reported, along with a further 17 cases with at least five deaths in 2022 [up to 16 March 2022 (*11*)]. These recently increased numbers may be due to enhanced respiratory and pneumonia surveillance during the COVID-19 pandemic, but the situation needs to be closely monitored. Reassuringly, no onward human-to-human transmission has been detected from any of these H5N6 or H5N8 HPAIv infections. To date, only a single human case with the recent 2.3.4.4 H5N1 HPAIv has been reported—a 79-year-old man in the United Kingdom who was infected from local ducks in mid-December 2021 but who remained asymptomatic (*12*).

What would raise the concern for H5N1 HPAIv for humans? A recent European assessment (*13*) determined that the risk of infection with H5Nx for the general population in Europe was low, and for occupationally exposed people, the risk was low to medium (but with high uncertainty) (*14*). Further adaption would need to occur in this current virus to increase its ability to transmit between humans efficiently, a process that is only partially understood; although this has not occurred yet, the persistence of H5Nx HPAIv increases the likelihood. As with COVID-19, if this adaption occurred and the virus became efficient at airborne transmission between humans, it would be virtually impossible to control.

What can be done to curb the H5N1 HPAIv threat? Although H5N1 and H5Nx HPAIv are, at present, a low threat to humans, it is early in the outbreak and each potential pandemic threat should be taken seriously. To gauge the risk, international organizations such as WHO, OIE, and the Food and Agriculture Organization of the United Nations (FAO) all have programs to investigate HPAIv in humans, poultry, and wild birds. Furthermore, continued investment in surveillance of wild birds and poultry and of humans at the human-poultry interface are critical. To mitigate the potential pandemic threat from H5Nx viruses, WHO continually prepares a stock of those viruses (and others) that are considered to be potentially dangerous to humans, which can be rapidly accessed for vaccine manufacture if required. Effective H5 HPAIv vaccines are also important in protecting poultry, and their use will reduce spillover into wild birds and humans. In 2017, mass poultry vaccination with a combined H5 and H7 vaccine was successful in controlling H7N9 poultry outbreaks and human infections in China. After mass vaccination and the closure of live-bird markets, H7N9 human infections dropped from 759 (with 281 deaths) in 2017 to only two cases in 2018 (*15*). Notably, countries that do not use poultry influenza vaccines and instead rely on the culling of infected birds to control HPAIv outbreaks may now consider the implementation of a vaccination program, given the impact of the current H5N1 outbreak. Additionally, measures such as reduction of flock size and density and avoidance of poultry production in waterbird-rich areas have been proposed to prevent spillovers of HPAIv into wild birds. The ongoing 2021–2022 wave of avian influenza H5N1 is unprecedented in its rapid spread and extremely high frequency of outbreaks in poultry and wild birds and is a continuing potential threat to humans. ■

Spread of H5Nx avian influenza

The H5Nx (2.3.4.4) highly pathogenic avian influenza viruses (HPAIv) have spread globally since 2014, with unprecedented outbreaks of H5N1 infections in poultry and wild birds in 2021 and 2022, with some spillover human infections. Wild-bird migratory flyways can be linked with the occurrence of these outbreaks (arrows).



The maps are simplified representations of avian HPAIv lineage 2.3.4.4 outbreaks from the Emergency Prevention System for Animal Health (EMPRES) of the Food and Agriculture Organization of the United Nations. See text for references to human cases.

GRAPHIC: M. WILLE AND I. G. BARR

REFERENCES AND NOTES

1. Animal and Plant Health Inspection Service, US Department of Agriculture; https://bit.ly/36BesKO.
2. OIE, "Avian influenza" (OIE, 2016); https://bit.ly/3qp9xE1.
3. FAO, "H5N1 HPAI Global overview: January–March 2012" (FAO, 2012); https://bit.ly/3L2TGmd.
4. V. Caliendo *et al.*, bioRxiv 2022.01.13.476155 (2022).
5. Wageningen University and Research, "Bird flu detected in dead Knots washed up on the Wadden Sea," 24 December 2021; https://bit.ly/3Ex3qTP.
6. ISID, ProMED; http://www.promedmail.org/post/8700909.
7. ISID ProMED; http://www.promedmail.org/post/8700696.
8. FAO, Global AIV with zoonotic potential, 23 March 2022; https://bit.ly/3L4x0C2.
9. Animal and Plant Health Agency, "Updated outbreak assessment #9" (VITT/1200 HPAI, 2022); https://bit.ly/3irbfQL.
10. O.G. Pyankova *et al.*, *Euro Surveill.* **26**, 2100439 (2021).
11. BNQ News, "Tracking human cases of H5N6 bird flu," *BNQ News*, 28 October 2021; https://bit.ly/37QtFIE.
12. I. Oliver *et al.*, *Euro Surveill.* **27**, 2200061 (2022).
13. European Food Safety Authority *et al.*, *EFSA* **19**, e07108 (2021).
14. Centers for Disease Control and Prevention; https://bit.ly/37EZGTO.
15. C. Li *et al.*, *Cold Spring Harb. Perspect. Med.* **11**, a0383492021 (2020).

Published online 26 April 2022
10.1126/science.abo1232

RETROSPECTIVE

Eugene N. Parker (1927–2022)

Pioneer of theoretical solar and plasma astrophysics

By **Stuart D. Bale**

Astrophysicist Eugene N. ("Gene") Parker, who led the field of solar physics into the space age, died on 15 March. He was 94. Gene was first to identify many of the fundamental concepts of Solar System plasma physics, including the physics of solar wind acceleration, coronal heating and magnetic reconnection, cosmic ray diffusion, and a conceptual structure of the heliosphere and its interaction with the interstellar medium. His foundational work on the theory of the solar wind led to the flagship NASA mission that bears his name: the Parker Solar Probe.

Gene was born in Houghton, Michigan, in 1927 and attended Michigan State University, where he received his BS in physics in 1948. He earned his PhD at the California Institute of Technology (Caltech) in 1951 working with physicist Leverett Davis Jr. on problems of the interaction of dust with the interstellar medium. After a brief appointment at the University of Utah, Gene joined the faculty at the University of Chicago, where he remained until his retirement in 1995.

Gene's early career coincided with the beginning of the space age. At that time, the prevailing wisdom was that interplanetary space was either a vacuum with sporadic solar perturbations or a solar atmosphere in hydrostatic equilibrium. However, subtle clues suggested something quite different. German astronomer Ludwig Biermann had made a convincing case that one component of a bifurcated comet tail always pointed away from the Sun, guided by a hypothetical "solar corpuscular radiation," whereas the other part was dominated by photon pressure and Kepler's laws.

Gene recognized that Biermann's corpuscular radiation was consistent with a hydrodynamic, transonic, escaping atmosphere that he called the "solar wind." His controversial 1958 solar wind theory, a kind of velocity-space analog of a rocket engine nozzle, was published only after the editor of the *Astrophysical Journal* overruled the reviewers. However, the theory provided predictions for measurements near Earth and suggested the existence of a "heliosphere," a plasma bubble surrounding the Sun with a discontinuous interface to the interstellar medium. The general framework of the model was verified by Soviet and American spacecraft measurements in one of the early scientific triumphs of the space age.

A 1960 Space Science Board report (the so-called Simpson Report) recommended both a solar probe to investigate the origins of Gene's solar wind and an outer Solar System probe to study the interaction of the heliosphere with the interstellar medium. Just 2 years after the founding of NASA, Gene's work was already setting the goalposts for major space mission objectives.

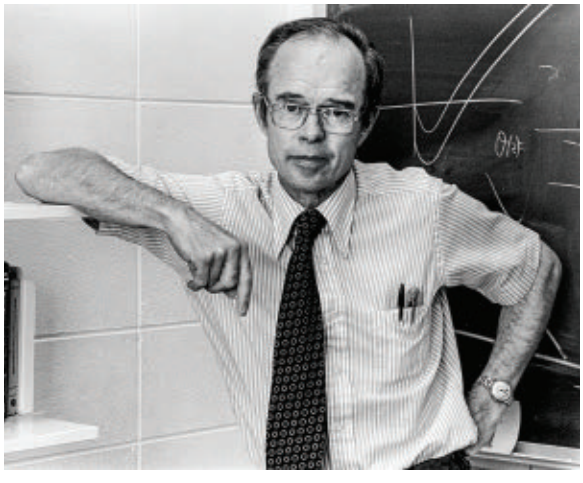


PHOTO: UNIVERSITY OF CHICAGO

Gene went on to define, and often resolve, many of the major problems of solar astrophysics. He proposed the concept of magnetic buoyancy to describe the origin and dynamics of sunspots, and he pioneered the models of cosmic ray diffusion and transport. He made major contributions to the theory of magnetic reconnection, the process that triggers solar flares and magnetospheric substorms. He elucidated the nature of the Maxwell stress tensor, showing that the lowest energy states are sheets of plasma currents. The nanoflare model that Gene proposed to explain solar coronal heating is still used today. In a major contribution to turbulent dynamo theory, he showed that a poloidal magnetic field arises directly from the dynamics of a toroidal field, an insight that completed the alpha-omega dynamo theory that explains the origin of planetary and astrophysical magnetic fields. Gene wrote four books on topics of astrophysical plasmas, each in his trademark dialectical style.

In 2010, NASA selected my team to design and build scientific instrumentation for the mission to investigate the origin and acceleration of the solar wind. The solar probe of the 1960 Simpson Report was finally coming to fruition. I asked Gene to join our team as a senior adviser, and he happily accepted the role. "My advice is free and worth every penny of that," he quipped. As we waded through theories to define our measurement requirements, he urged us to keep it simple. He stood up at one of our early meetings to remind us thoughtfully that "hydrodynamics and magnetohydrodynamics are fundamental, regardless of popular opinion."

Our Solar Probe Plus spacecraft—renamed the Parker Solar Probe in Gene's honor—launched in August 2018. It was the first time that a NASA spacecraft had carried the name of a living scientist. The Parker Solar Probe spacecraft has so far made 11 orbits of the Sun, diving to within 13 solar radii of the solar photosphere. Gene stayed abreast of the Parker Probe results as we sent him preprints and publications. As his health declined, he was content to cheer us on from the sidelines with the occasional handwritten letter. Many of Gene's predictions and theories are being borne out by Parker Solar Probe measurements, notably the existence of an underlying hydrodynamic wind speed profile—popular opinion be damned.

Gene was known as a supportive and inspiring mentor to his students. He was deeply committed to striking a good work-life balance; his former student Bill Collins recalls him saying, "The University of Chicago pays me for 40 hours per week and that's what they're going to get." Clearly, he put that time to good use: His many honors include the US National Medal of Science, the Crafoord Prize awarded by the Royal Swedish Academy of Sciences, the Kyoto Prize, the Arctowski Medal of the US National Academy of Sciences, and the American Physical Society Medal for Exceptional Achievement in Research.

The history of modern solar physics followed the arc of Gene's career, and his death is an opportunity to reflect on both his achievements and the state of the field. Solar and heliospheric physics are now entering a golden age, with extraordinary commitments from NASA, the US National Science Foundation, and the European Space Agency. Gene Parker's many insights played a fundamental role in making that possible. ■

Department of Physics and Space Sciences Laboratory, University of California, Berkeley, CA, USA.
Email: bale@berkeley.edu

10.1126/science.abq3164

POLICY FORUM

GLOBAL HEALTH

A global system for the next generation of vaccines

COVID-19 has shown that hurdles can be overcome

By **Nimalan Arinaminpathy[1,2], Chadi M. Saad-Roy[3], Qiqi Yang[4], Isa Ahmad[1,2], Prashant Yadav[5,6], Bryan Grenfell[4,7]**

The rapid development of vaccines against COVID-19 raises questions about what could be achieved in vaccines to other major diseases. Influenza presents an important case study; it is one of the few infections that causes substantial public health burden in its endemic form while also having proven pandemic potential. We offer a first step in bringing together the value proposition of future influenza vaccines considering two key characteristics: the breadth of protection that vaccines offer (against individual strains, all strains within a subtype, multiple subtypes, or all subtypes) and the duration for which protection remains effective (see the figure). We examine implications of these characteristics, from both epidemiologic and economic perspectives, and discuss how a future market for influenza vaccines might best align public health and economic incentives. Although many of these factors are specific to influenza, we consider comparisons with other major infections, such as tuberculosis and COVID-19.

Current influenza vaccines, mainly targeting the variable region of the haemagglutinin surface glycoprotein of the virus (*1*), cannot be used to protect a population in advance of a pandemic and are only deployed in routine immunization against specific variants of seasonal influenza. By contrast, through offering broad protection, new vaccines could mitigate equally against both seasonal and pandemic influenza. Such vaccines would be distinct among immunization programs by having value in both routine and strategic modes of deployment. Moreover, current influenza vaccines are mostly still produced by using the same egg-based methods that have been used for roughly the past 70 years. New vaccines would transform these production methods. For example, both mRNA and viral vector technology bypass the need for candidate vaccine viruses to initiate the production process, requiring only the genetic sequence of the target antigens. Although such vaccines will depend on new technologies for their manufacture, the unprecedented pace and scale of rollout of COVID-19 vaccines have shown that such hurdles can be overcome.

EPIDEMIOLOGICAL IMPLICATIONS

The basic reproduction number for seasonal and pandemic influenza (R_0)—the expected number of secondary cases produced by a typical infectious individual in a fully susceptible population—is typically less than 2, meaning that even partial vaccination coverage and efficacy can have strong effects in reducing the burden of influenza. It is likely that a future influenza vaccine would offer less than lifelong protection, necessitating repeated vaccination over a lifetime. Depending on the duration of immunity, repeat vaccination may not need to be as frequent as every year. If the vaccine targets "conserved" antigens that do not evolve as rapidly as the variable region of haemagglutinin (the target of current vaccines), the antigenic formulation of the vaccine may need not be updated as often as with current seasonal vaccines.

Previous modeling work has illustrated that if the duration of immunity and the frequency of immunization are sufficiently high that population immunity is allowed to accumulate over time, a future influenza vaccine could exert greater reductions in seasonal influenza burden than that of a conventional vaccine with the same coverage and efficacy (*2*). Multiyear cohort studies of vaccinated versus unvaccinated individuals would provide valuable data on the duration of vaccine-induced protection that will be important in planning future vaccination programs.

On breadth, although an ideal vaccine would be effective against all subtypes of influenza A, in practice it has proven challenging to achieve such universal breadth of protection. Previous work has shown how infection-induced, T cell–mediated heterosubtypic immunity could protect against severe disease from pandemic influenza in humans (*3*). Several vaccine candidates aim to raise T cell immunity against targets such as the nucleoprotein and the matrix protein [for example, (*4*)], which are conserved across influenza subtypes.

On the other hand, recent work has focused on antibody-mediated immunity against the haemagglutinin stalk, which, being critical for the stability of the overall haemagglutinin molecule, shows far less variability than the "head" region [for example, (*5*)]. However, stalk-based immunity is not universal: Because H1 and H3, the two subtypes of influenza A currently causing seasonal influenza in humans, belong to different clades, stalk-reactive antibodies to H1 may offer only limited cross-protection to H3, and vice versa. Vaccinating against H3 may therefore open an ecological niche, allowing greater epidemics of H1 than currently occur, and vice versa. Analogous effects have been observed for pneumococcus, in which widespread vaccination uptake against a limited set of serotypes has typically been followed by an increase in carriage of nonvaccine serotypes (although not necessarily accompanied by increases in disease) (*6*).

Overall, for any future influenza vaccine offering less-than-universal protection, it will be critical to adopt complementary vaccination strategies to close the "protection gap" (see the figure). For example, a future vaccination program could combine a stalk-based H3 vaccine with a conventional H1 vaccine, or a combination of stalk-based H1 and H3 vaccines.

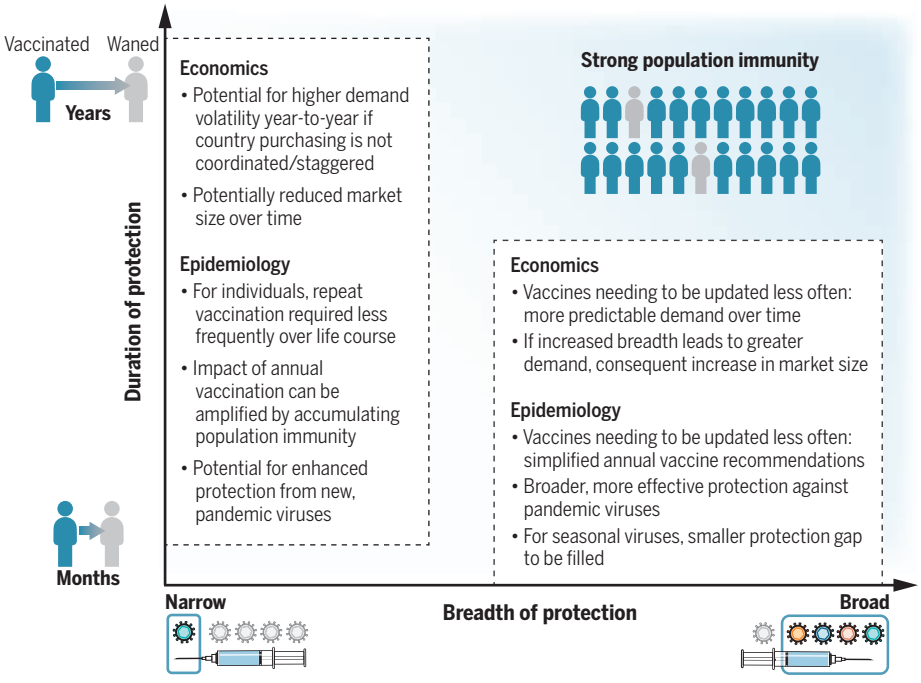
ECONOMIC IMPLICATIONS

Although improved breadth and duration of protection both have positive implications for epidemiological impact, they could act in opposing directions in the context of manufacturing and distribution (see the figure). A future influenza vaccine that offers broad protection could have higher global demand than that of current influenza vaccines, giving rise to a substantially expanded market for manufacturers to serve. If such a vaccine does not need to be updated as frequently as current vaccines, as a result of targeting more conserved influenza anti-

[1]MRC Centre for Global Infectious Disease Analysis, Department of Infectious Disease Epidemiology, School of Public Health, Imperial College London, London, UK. [2]Abdul Latif Jameel Institute for Disease and Emergency Analytics, School of Public Health, Imperial College London, London, UK. [3]Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ, USA. [4]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA. [5]Technology and Operations Management, INSEAD, Fontainebleau, France. [6]Center for Global Development, Washington, DC, USA. [7]Princeton School of Public and International Affairs, Princeton University, Princeton, NJ, USA. Email: nim.pathy@imperial.ac.uk

Economic and epidemiologic considerations

Boxes show key factors associated with improving vaccine performance in terms of duration (vertical axis) and breadth of protection (horizontal axis).


Vaccinated
Waned
Years
Duration of protection
Economics
• Potential for higher demand volatility year-to-year if country purchasing is not coordinated/staggered
• Potentially reduced market size over time
Epidemiology
• For individuals, repeat vaccination required less frequently over life course
• Impact of annual vaccination can be amplified by accumulating population immunity
• Potential for enhanced protection from new, pandemic viruses
Strong population immunity
Economics
• Vaccines needing to be updated less often: more predictable demand over time
• If increased breadth leads to greater demand, consequent increase in market size
Epidemiology
• Vaccines needing to be updated less often: simplified annual vaccine recommendations
• Broader, more effective protection against pandemic viruses
• For seasonal viruses, smaller protection gap to be filled
Months
Narrow
Breadth of protection
Broad

GRAPHIC: N. CARY/SCIENCE

gens than haemagglutinin, this would also simplify the manufacturing process (*7*).

However, although longer-duration immunity could have important public health benefits, it could also pose important considerations for manufacturers. For example, comparing a vaccine that needs to be administered every year against one that is only required every 5 years, over time the former would have five times the potential market as the latter. This simple comparison could be moderated by any increase in demand associated with longer-duration immunity; understanding such effects will be important in planning for manufacturing and distribution of future vaccines.

Vaccines conferring longer-duration immunity will also require careful planning in their manufacturing and deployment. As an extreme but illustrative example, if a country's immunization program involves mass vaccination every 5 years, this would naturally lead to boom-and-bust cycles in demand; in intervening years, manufacturers will need to repurpose production lines for other vaccines. Where different countries follow different immunization programs, this could lead to demand volatility, potentially compromising sustainability for manufacturers. Alternative immunization schedules could help to smooth demand over time, such as "cohort" vaccination strategies focused each year on 5-year-olds, 10-year-olds, and so on. Such strategies are routinely used for childhood vaccines such as DTaP against diphtheria, tetanus, and pertussis.

Other approaches to address demand volatility could include centralized purchasing mechanisms to coordinate and stagger vaccine procurement among different countries, helping to create a market that remains stable over time. Examples of such mechanisms include the Pan American Health Organization (PAHO) revolving fund and the United Nations Children's Fund (UNICEF) Supply Division, which acts on behalf of Gavi, the Vaccine Alliance.

GLOBAL PRODUCTION CAPACITY

At present, production of influenza vaccines is heavily concentrated in high-income settings. Manufacturers in high-income countries, accounting only for 16% of the global population, account for an estimated 69% of seasonal influenza vaccine production and 80% of potential pandemic vaccine production (*8*). These disparities mirror global imbalances in the uptake of influenza vaccines: Although the United States has among the world's largest coverage of routine influenza vaccination as a proportion of its population, uptake is much more limited in India and other low- and middle-income countries (LMICs). Reasons for such low uptake (*9*) include lack of awareness of the disease burden of influenza, as well as the need to vaccinate annually, and the cumbersome process of deploying vaccines that must be reformulated regularly as the virus evolves. It is likely that a future, broadly protective vaccine, with similar efficacy to that of current vaccines but offering multiyear protection while not needing to be reformulated so often, would address some key issues for uptake in LMICs.

Alongside increasing uptake to facilitate the emergence of a global market, we argue that there will be several benefits to promoting regional production to meet local needs while leveling current imbalances in global production capacity. From an epidemiological perspective, future influenza vaccines could be used strategically for pandemic preparedness and response. As demonstrated during the COVID-19 pandemic, global supply of pandemic vaccination is not always aligned with global need because countries prioritize coverage among their own populations. If increased future demand for new influenza vaccines facilitates the establishment of regional production capacity to meet local needs, this distributed capacity will be valuable in ensuring that all countries have the necessary access to vaccines for pandemic response. This benefit need not be limited to influenza pandemics, with new technologies such as mRNA and viral vectors being applicable to other infections as well. As long as regional vaccine manufacturing sites are sufficiently flexible, they could be deployed rapidly against a wide array of emerging infections (*10*). New initiatives already underway to expand manufacturing capacity for COVID-19 and other vaccines in Africa and elsewhere could represent important steps toward equitability in global pandemic preparedness.

From an economic perspective, for both seasonal and pandemic vaccination, an expanded market with the right number of manufacturers allows production to benefit from economies of scale (*11*), potentially allowing lower vaccine prices at least for the poorest countries. A large, globally distributed market could establish a virtuous cycle by reducing local costs of broadly protective vaccines, which in turn could facilitate increased uptake (and consequently commercial attractiveness). Gavi exerted such effects on the global market for the pentavalent vaccine: Their coordinated market-shaping initiatives led to more manufacturers entering the market, which in turn reduced vaccine prices (*12*).

A much-expanded market could also lower costs by facilitating entry of new manufacturers. Currently, only a few low- or lower-middle-income countries have capacity to produce vaccines that meet stringent quality standards [approved by stringent regulatory authorities or being World Health Organization (WHO)–

prequalified]. Use of newer platforms such as recombinant proteins, adenovirus vectored, or lipid-encased mRNAs is largely limited to India and China (*13*). Developing such manufacturing capacity in other countries requires intensive "soft" infrastructure in the form of chemistry, manufacturing, and control (CMC) staff and "hard" infrastructure in the form of new equipment for manufacturing and laboratory and analytical testing. New manufacturers who invest in this infrastructure are likely to operate on a smaller scale compared with that of manufacturers in India and China and are therefore likely to see increased costs of production. A much expanded market could therefore facilitate the entry of these manufacturers by offsetting these increased costs.

ROLES FOR DONORS AND POLICY-MAKERS

Catalyzing the development of such a global market is likely to require effective partnerships between industry, policy-makers, and academia. There will be a critical role for international agencies and donors to play, not only in the research and development of new vaccines but also in their manufacturing and distribution.

Mechanisms such as advance market commitments could attract greater industry participation in the process of vaccine development. To the extent that new vaccines depend on new technologies, they will also require new mechanisms for evaluation of their performance. Pathways for regulatory approval will need to be updated, to incorporate these new mechanisms. A global market will require updates to regulatory approval to be coordinated at a global level to facilitate the deployment of new vaccines worldwide. It is likely that licensure of future vaccines will proceed in steps rather than leaps—for example, with breadth and duration of protection being steadily improved with each successive generation of vaccine development. Therefore, although "target product profiles" that identify ideal characteristics for future vaccines would offer helpful reference points in vaccine development and international regulatory coordination, it will be equally important for donors and international agencies to plan how best to recognize and reward intermediate stages toward achieving those ideals.

Once vaccines are developed and licensed, there will be a need for further interventions in their distribution, both on the demand side and the supply side. Creating a sustainable demand side requires mechanisms for long-term public funding of vaccine purchases by country governments, global coordination of influenza vaccines for LMICs, and if needed, using pooled procurement structures such as UNICEF and the PAHO revolving fund for LMIC purchase of influenza vaccines. On the supply side, there will be a need for concessional loans and other forms of investment from development finance institutions (such as the International Finance Corporation) to finance the capital expenditure for setting up new sites in LMICs for influenza vaccine manufacturing.

In addition to plant and equipment and technical know-how, vaccine manufacturing—especially using newer cell-based platforms and mRNA—requires many complex input components, some of which have limited sources of supply. A globally coordinated mechanism is needed to manage the material, know-how, and other critical resources required for the global market to operate. The WHO Global Action Programme initiative offers an example of such a coordinated mechanism, in the context of current influenza vaccines (*8*). As new technology platforms emerge, new coordination architectures may have to be designed to accommodate their deployment.

COMPARISONS WITH OTHER INFECTIONS

To illustrate how the value propositions for future vaccines for other diseases could depend on the disease in question, we consider two other respiratory diseases, COVID-19 and tuberculosis (TB).

Like influenza, coronaviruses are established in the human population and have proven pandemic potential. Although licensed vaccines to SARS-CoV-2) raise immunity to the surface spike protein, there is increasing interest in targeting more conserved antigens to raise more broadly protective immunity. Although it will be challenging to achieve a truly universal coronavirus vaccine, recent work has shown promising signs of raising broad protection against sarbecoviruses [for example, (*14*)], a lineage to which SARS-CoV-2 belongs, along with SARS-CoV-1, which first emerged in humans in 2003. Owing to the global burden of SARS-CoV-2, and the need to protect against future coronavirus pandemics, many of the considerations raised here (see the figure) would apply to future coronavirus vaccines as well. However, endemic coronaviruses have only modest contribution to the burden of respiratory illness, particularly in comparison with seasonal influenza; SARS-CoV-2 may, over time, similarly transition to being an endemic virus causing mostly mild illness. Accordingly, compared with influenza vaccines, the public health value proposition of future coronavirus vaccines may have a greater emphasis on pandemic protection than on routine immunization. However, a sizable burden of "long covid" after endemic SARS-COV-2 infection could alter this balance.

For TB, the world's leading cause of death from infectious disease, the public health case for new, effective vaccines is clear. However, with global TB burden overwhelmingly concentrated in LMICs, the potential market for future TB vaccines is not as commercially promising as for other vaccines. Hence, despite a recent phase 2b trial of the M72/AS01e vaccine showing promising signs of efficacy (*15*), it has been challenging to find industry partners who are willing and capable of taking this candidate forward to phase 3 trials. The Bill and Melinda Gates Medical Research Institute (Gates MRI) will support clinical trials of the vaccine, having recently acquired the license from Glaxo SmithKline. Even though future vaccines against TB and other diseases of poverty may well benefit from new technologies deployed against COVID-19, international public health agencies and donors will need to play an even more active role than in the context of influenza vaccines.

A globally distributed market for influenza vaccines, supported by regional production capacity to meet regional need, would be one way of aligning public health, supply resilience, and commercial interests but will not necessarily develop on its own; international public health agencies and donors will have a key role to play, in catalyzing the creation of such a truly global system. ■

REFERENCES AND NOTES

1. M. T. Osterholm, N. S. Kelley, A. Sommer, E. A. Belongia, *Lancet Infect. Dis.* **12**, 36 (2012).
2. R. Subramanian, A. L. Graham, B. T. Grenfell, N. Arinaminpathy, *PLOS Comput. Biol.* **12**, e1005204 (2016).
3. S. Sridhar *et al.*, *Nat. Med.* **19**, 1305 (2013).
4. R. D. Antrobus *et al.*, *Mol. Ther.* **22**, 668 (2014).
5. R. Nachbagauer *et al.*, *Nat. Med.* **27**, 106 (2021).
6. D. M. Weinberger, R. Malley, M. Lipsitch, *Lancet* **378**, 1962 (2011).
7. L. J. Kornish, R. L. Keeney, *Oper. Res.* **56**, 527 (2008).
8. E. Sparrow *et al.*, *Vaccine* **39**, 512 (2021).
9. J. R. Ortiz, K. M. Neuzil, *J. Infect. Dis.* **219** (Suppl_1), S97 (2019).
10. P. Yadav, A. Desir, *INSEAD Knowl.* 3 November 2021; https://knowledge.insead.edu/operations/boosting-vaccine-production-needs-the-right-degree-of-flexibility-17621.
11. S. Plotkin, J. M. Robinson, G. Cunningham, R. Iqbal, S. Larsen, *Vaccine* **35**, 4064 (2017).
12. M. Malhame *et al.*, *Vaccine* **X 2**, 100033 (2019).
13. National Academies of Sciences, Engineering, and Medicine, National Academy of Medicine, *Globally Resilient Supply Chains for Seasonal and Pandemic Influenza Vaccines* (The National Academies Press, 2021).
14. D. R. Martinez *et al.*, *Science* **373**, 991 (2021).
15. D. R. Tait *et al.*, *N. Engl. J. Med.* **381**, 2429 (2019).

ACKNOWLEDGMENTS

The authors thank S. Suvanand and D. Lewinsohn for helpful comments on future TB vaccines. In addition to the funding for this work from FluLab, C.M.S.-R. acknowledges support from a Postgraduate Doctoral Scholarship of the Natural Sciences and Engineering Research Council of Canada and from a Charlotte Elizabeth Procter Fellowship of Princeton University. N.A. acknowledges support from the MRC Centre for Global Infectious Disease Analysis (reference MR/R015600/1).

10.1126/science.abm8894

CALL FOR PAPERS

SPJ
SCIENCE PARTNER JOURNALS

北京理工大學
BEIJING INSTITUTE OF TECHNOLOGY

中国航天
中国空间技术研究院
China Academy of Space Technology

Space: Science & Technology

Space: Science & Technology is an online-only, Open Access journal published in affiliation with **Beijing Institute of Technology (BIT)** and distributed by the **American Association for the Advancement of Science (AAAS)**. BIT cooperates with China Academy of Space Technology (CAST) in managing the journal. The journal promotes the exploration and research of space worldwide, to lead the rapid integration and technological breakthroughs of interdisciplinary sciences in the space field, and to build a high-level academic platform for discussion, cooperation, technological progress and information dissemination among professional researchers, engineers, scientists and scholars.

Submit your research to *Space: Science & Technology* today!

Learn more at **spj.sciencemag.org/space**

The Science Partner Journal (SPJ) program was established by the American Association for the Advancement of Science (AAAS), the nonprofit publisher of the *Science* family of journals. The SPJ program features high-quality, online-only, Open Access publications produced in collaboration with international research institutions, foundations, funders and societies. Through these collaborations, AAAS furthers its mission to communicate science broadly and for the benefit of all people by providing top-tier international research organizations with the technology, visibility, and publishing expertise that AAAS is uniquely positioned to offer as the world's largest general science membership society.

Visit us at **spj.sciencemag.org**

@SpaceSciTech @SPJournals

OPEN ACCESS

ARTICLE PROCESSING CHARGES WAIVED UNTIL JULY 2023

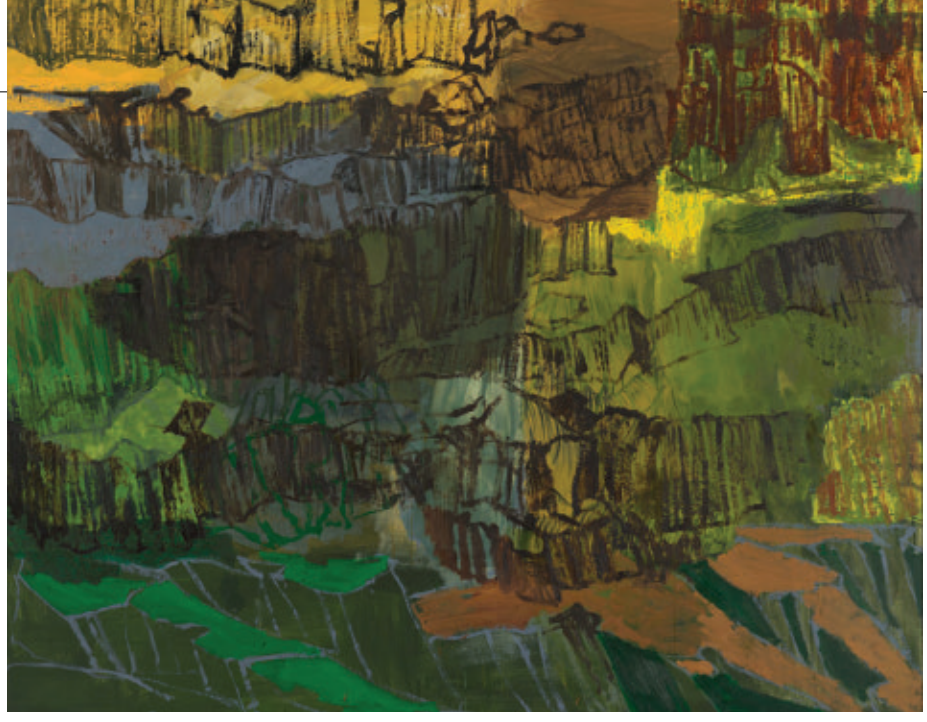
BOOKS *et al.*

GEOLOGY

Landscapes of the Anthropocene

The late artist Per Kirkeby's preoccupation with geology is on display in a new exhibition

By **Deborah Dixon**



Geological cross sections come to life in *Geologische Nachrichten (Geological Messages)*, painted in 1999.

Geological Messages, the collection on show at the Michael Werner Gallery in London, consists of 24 paintings in various mediums by the acclaimed Danish artist Per Kirkeby (1938–2018). Over the years, reviewers have compared the pieces created by Kirkeby—who initially trained as a geologist—with various artistic traditions in order to draw out how he deals with the materialities of light, shadow, substance, and system. But this collection asks us to consider Kirkeby's "decades-long preoccupation with the world as geologically unstable, ductile, and everywhere moving between states of mineral complexity and processes of collapse" (*1*).

Geologists have deployed a wide array of visualization techniques, from sketches and lithographs to models and scans, to investigate the transformation of landforms. Such techniques animate a series of aesthetics specific to geology and include a mapping of surface–subterranean relationships, the forensic examination of grain-scale characteristics for signs of cosmic import, and the slice through deep time that is the geological cross section.

The earlier works on display in this exhibition are mixed media on Masonite, a material formed by steaming, stretching, and pressing wood fibers. We see in these paintings the exposure of the subterranean in striated sandstone buttes eroded by water and wind, and in caves dissolved into karst by rainwater percolating downward.

Geological Messages: Paintings from 1965–2015
Per Kirkeby
Michael Werner Gallery, London, UK, through 21 May 2022

This focus on the exposure of the depths of Earth has a long geological lineage. Kirkeby's 1967 painting *Dunkle Höhle (the Dream about Uxmal and the Unknown Grottos of Yucatan)*, which looks outward from a limestone grotto, echoes, for example, the sketched excavations of Kirkland Cave in Yorkshire by Victorian geologist William Buckland (*2*), whose metaphysics of sequential extinctions and emergences veered between creationism and catastrophism.

In Kirkeby's 1992 oil on canvas painting *Inferno IV*, we can find another geological aesthetic. In place of the tension between surface form and depth, there appear granularities and fractures. This piece echoes the thin-section microscopies of minerals, ores, and meteorites where the passage of white light through the material reveals various hues and intensities, as well as a textural world of shapes, cleavages, facets, and inclusions. Peering at Kirkeby's forensic grammar of scraped-on paint and scribbled lines, we can find a narration of tectonic uplift, eruptions, the appearance and disappearance of oceans, and the continual dynamism not just of Earth but of the cosmos within which it revolves.

In *Geologische Nachrichten (Geological Messages)*, an oil on canvas painted in 1999, it is the geological cross section that comes into view. Planes and folds are mapped out in block color and serried black outlines. But this is a cross section that refuses the careful stratigraphy of a Charles Lyell, James Hutton, or G. K. Gilbert. Kirkeby's patchwork terrains seem to exist in their own separate times and spaces, their planar surfaces illuminated by different light sources.

This strangeness unfolds in Kirkeby's more recent, untitled works, which retain lithic components but situate them amid organic formations such as tree trunks, the grains of which echo fault lines. Across these visually lush paintings there are branches and leaves, as well as a snake, a rabbit, and birds. The mixed media on Masonite of these paintings reiterates the mix of shapes, colors, and lines that are uncannily suggestive of undergrowth, crystals, swarms, and the fronds of underwater—or aerial—life-forms.

As rock becomes part of the Anthropocene's hybrid, strange ecologies, a different geological message emerges. The Arctic and northern European landscapes that Kirkeby references are undergoing ever-quickening transformations under anthropogenic climate change, including the movement of mass materials. Communities that for generations have known the alignments of rock, vegetation, rivers, and animal life are facing increasing uncertainty.

Kirkeby's "geological messages" are a very welcome reminder of the shared aesthetics and practices that have animated geology and art, as well as the planetary and cosmic narratives that geology gives us. For audiences now, these messages seem to be telling the time of the Anthropocene, not as another stratum that overlays an earthly archive built from stone but as a condition of estrangement and alienation that requires new ways of connecting with Earth. ■

The reviewer is at the School of Geographical and Earth Sciences, University of Glasgow, Glasgow G12 8QQ, UK. Email: deborah.dixon@glasgow.ac.uk

REFERENCES AND NOTES

1. A. Smith, Aust. *N. Z. J. Art* **12**, 178 (2012).
2. W. Buckland, *Phil. Trans. R. Soc.* **112**, 171 (1822).

10.1126/science.abp9703

IMAGE: COURTESY OF MICHAEL WERNER GALLERY, NEW YORK AND LONDON

Rethinking the "Western" revolution in science

A historian sees global cultural and geopolitical roots in Europe's scientific breakthroughs

By **Jorge Cañizares-Esguerra**

Historian James Poskett sees modern science as the product of ceaseless cross-cultural, geopolitical transformations. In his new book, *Horizons*, Poskett argues that the so-called "Western" revolution in science was actually the result of global transcultural and transregional interactions, all triggered by global socio-cum-geopolitical shifts. He brings his model of transculturation and global geopolitics transformations to bear on nearly every major breakthrough in modern science, from Newton's physics to Lavoisier's chemistry to Mendel's genetics. The result is a bold new interpretation of the history of the field.

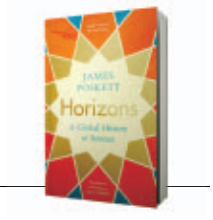
Take, for example, the work of Copernicus. Poskett locates Renaissance astronomy within the much larger context of the expansion of Islam in North Africa and Eurasia. Islam, he reveals, demanded great accuracy in the observations of solar, lunar, and planetary movements, which were used to organize ritual calendars and sanction political power. The fall of Constantinople in 1453 triggered an influx of texts, exiles, and learned captives into Europe, and Copernicus drew on these resources when formulating his heliocentric theory of the Universe. It was therefore, in Poskett's estimation, the combined effort of many scholars, rather than the work of a lone genius, that led to the demise of Aristotelian and Ptolemaic geocentric models.

Similarly, Poskett demonstrates how all the key evidence Isaac Newton relied on to revitalize physics came from comparative studies conducted in equatorial and Arctic locations. To reach isolated islands in the Pacific to obtain such data, nations needed considerable seafaring capabilities. Ultimately, Poskett argues, it was the Atlantic slave trade that made the accumulation of evidence for Newtonian physics possible.

According to Poskett, 19th-century industrialization, nationalism, settler colonialism, and imperialism drove the development of evolutionary biology, particularly Darwinian natural selection. The idea of evolution as the survival of the fittest was a trope that informed the development of national armies and frontier societies. Nineteenth-century Argentina, Russia, Japan, and China, he notes, excelled in paleontology and evolutionary biology.

It was the pursuit of communication in scattered imperial polities that led to breakthroughs in telegraphy and radio, particularly in Russia, Japan, and China, argues Poskett. The rush to industrialization in the 19th century, in turn, sparked much research in chemistry. During this period, the Russian chemist Dmitri Mendeleev spearheaded inquiries into the periodic table, and the Japanese physicist Hantaro Nagaoka headed up efforts that led to the first model of the atom.

Horizons: The Global Origins of Modern Science
James Poskett
Mariner Books, 2022.
464 pp.



Poskett sees Marxist ideologies and Cold War conflict as two of the main forces behind the global spread of relativity, quantum mechanics, and genetics. The new physics advanced by Einstein and Bohr was received by Bolshevik radicals in Russia as a welcome challenge to bourgeois complacency, he argues. Nationalists and republican radicals in India and China also embraced these ideas with gusto. Meanwhile, the fear of famine and atomic warfare during the Cold War greatly facilitated the spread of population and agricultural genetics in Mexico, China, and India.

Poskett's main contribution with this book is to demonstrate that "European" knowledge has long been the result of global efforts and that science is intimately attached to colonialism, capitalism, slavery, industrialization, and geopolitical conflict. Poskett offers countless examples of non-European scientists whose research changed the sciences in radical new ways.

Yet, for all his sharp insights, Poskett remains firmly grounded in Eurocentric teleologies. In his account, science moves through familiar European markers of progress: Copernicanism, Newtonianism, Linnaean natural history, Maxwellian electromagnetism. And nearly all the non-European innovators he highlights first trained in either European or US institutions. Nor does he question the Eurocentric assumptions that underlie the historiographical categories that are traditionally used to organize the history of science (the print revolution, the Republic of Letters, the public sphere, the Enlightenment, democracy, and the Industrial Revolution). Despite these shortcomings, *Horizons* is nevertheless a challenging book that deserves a wide readership. ■

PHOTO: BABAK TAFRESHI/SCIENCE SOURCE

Islamic astronomy informed the work of Copernicus, argues Poskett.

The reviewer is at the Department of History, University of Texas at Austin, Austin, TX 78712, USA. Email: canizares-esguerra@austin.utexas.edu

10.1126/science.abo5229

LETTERS

Edited by **Jennifer Sills**

Retraction

We would like to retract the Report "Asymmetric remote C–H borylation of aliphatic amides and esters with a modular iridium catalyst" (*1*) after discovering that the reported enantioselective gamma-selective C–H borylation of *N*,*N*-dibenzylhexamide (**1a**) is not reproducible and that many of the nuclear magnetic resonance (NMR) spectra were manipulated. For example, the signal patterns of the ^{1}H NMR spectra of boronate (*R*)-**2b** (4-boryl-*N*-*tert*-butylhexamide) and (*R*)-**2g** (*tert*-butyl 4-boryl-hexanoate) (S129, S144) are identical, the baselines of the ^{13}C NMR spectra of (*R*)-**2b** and (*R*)-**2g** are identical (S130, S145), duplicative signals are found in the ^{1}H NMR spectrum of boronate (*R*)-**2f** (ethyl 4-boryl-hexanoate) (two triplets in 0.7 to 0.9 ppm and two multiplets in 2.2 to 4.1 ppm) (S141), and the baseline of (*R*)-**2f** (S142) is also identical to those of (*R*)-**2b** and (*R*)-**2g**. These issues undermine our confidence in the integrity of the study as a whole. We regret any confusion and apologize to the scientific community. All authors have agreed to retract this Report, and an institutional investigation of misconduct is underway.

Ronald L. Reyes[1,2], Miyu Sato[2], Tomohiro Iwai[2], Kimichi Suzuki[1,2], Satoshi Maeda[1,2], Masaya Sawamura[1,2]*

[1]Institute for Chemical Reaction Design and Discovery (WPI-ICReDD), Hokkaido University, Sapporo 001-0021, Japan. [2]Department of Chemistry, Faculty of Science, Hokkaido University, Sapporo 060-0810, Japan.
*Corresponding author.
Email: sawamura@sci.hokudai.ac.jp

REFERENCES AND NOTES

1. R. L. Reyes *et al.*, *Science* **369**, 970 (2020).

10.1126/science.abq4856

Evolutionary risks of osprey translocations

The osprey *Pandion haliaetus* is a mostly migratory, cosmopolitan raptor (*1*) that was historically extirpated or highly depleted in southern Europe. Because the species is charismatic, it has garnered public attention that has stimulated broad conservation action, fundraising, and political commitment. Osprey translocation programs have been a primary conservation strategy in Europe since 1996 (*1*). However, because evidence suggests that translocation programs that move individuals without accounting for local adaptations between populations have negative short- and long-term effects (*2*–*5*), this approach should no longer be used to meet conservation goals.

An osprey chick from an endangered population sits in a nest on the rocky cliffs of Corsica.

PHOTO: JEAN-MARIE DOMINICI

Translocations over the past several decades have contributed to increases in southern and western osprey populations (*1*). However, in most cases source populations were from distant geographical areas of central and northern Europe (*6*), from which native Mediterranean osprey populations differ in both genetics (*6*) and migratory behavior (*6*–*8*). Whereas native Mediterranean populations are residents or short-distance migrants within the Mediterranean basin (*7*), breeding birds from reintroduced populations in Spain and Portugal maintained a long-distance migratory pattern, similar to birds from the donor populations (*8*).

Reintroductions that are not consistent with the evolutionary history of a species, and that will likely alter the species' future evolution, could prove detrimental to conservation (*2*–*5*). Evidence suggests that mortality of translocated animals, mixing with neighboring local populations (*2*), and changes in migratory and dispersal behavior (*2*, *5*, *8*) are likely to take place in the short term. Long-term effects could include genetic homogenization that annihilates part of the genetic diversity of the species at broader spatial scales (*2*–*4*) and decreases in the species' ability to adapt to climate change.

In June, the local government of Generalitat Valenciana plans to release near Valencia, Spain, large numbers of ospreys originating in central Europe (*9*). Although the project was conceived as a way to improve conservation, it instead focuses on the short-term social and political benefits associated with the return of a totemic species. By prioritizing short-term goals, the plan disregards the demonstrated risks of such translocations and the recommendations of international conservation agencies (*10*–*12*).

We warn the scientific community and the Spanish authorities against the planned actions and recommend appropriate international assessments to ensure that all future reintroductions adhere to

science-based recommendations. Genetic and behavioral processes that occur on evolutionary time scales are at least as important as those that are readily visible at the time scales of most conservation actions (*10*–*12*).

Flavio Monti[1,2]*, Claudine Montgelard[3], Alexandre Robert[4], Andrea Sforzi[2], Rafel Triay[5], François Sarrazin[4], Olivier Duriez[3]
[1]Department of Physical Sciences, University of Siena, Earth and Environment, 53100 Siena, Italy. [2]Maremma Natural History Museum, 58100 Grosseto, Italy. [3]Center for Evolutionary and Functional Ecology, University of Montpellier, and Ecole Pratique des Hautes Etudes, Institut pour la Recherche et le Developpement, Centre National de la Recherche Scientifique (CNRS), Montpellier, France. [4]Centre d'Ecologie et des Sciences de la Conservation, Muséum National d'Histoire Naturelle, CNRS, Sorbonne Université, 75005 Paris, France. [5]Institut Menorquí d'Estudis, E-07702 Maó, Illes Balears, Spain.
*Corresponding author.
Email: flaviomonti00@gmail.com

REFERENCES AND NOTES

1. A. F. Poole, *Ospreys: The Revival of a Global Raptor* (Johns Hopkins University Press, Baltimore, MD, 2019).
2. C. B. Viverette, "Influence of historic landscapes and contemporary species management on Chesapeake Bay bald eagles and osprey," thesis, Virginia Commonwealth University, USA (2016).
3. M. Wheeler, "The genetics of conservation translocations: A comparison of North American golden eagles (*Aquila chrysaetos canadensis*) and bald eagles (*Haliaeetus leucocephalus*)," thesis, Duquesne University, USA (2014).
4. F. Jacobsen, M. Nesje, L. Bachmann, J. T. Lifjeld, *Conserv. Genet.* **9**, 581 (2008).
5. R. J. Burnside, C. Buchan, D. Salliss, N. J. Collar, P. M. Dolman, *R. Soc. Open Sci.* **7**, 200250 (2020).
6. F. Monti *et al.*, *Conserv. Genet.* **19**, 839 (2018).
7. F. Monti *et al.*, *J. Avian Biol.* 49, e01839 (2018).
8. B. Martín *et al.*, *Eur. J. Wildl. Res.* **65**, 72 (2019).
9. Informe Técnico, "Reintroducción del Águila Pescadora en la Comunitat Valenciana: Fase experimental 2019-2020" (Servei de Vida Silvestre, Direcció General de Medi Natural i d'Avaluació Ambiental, 2020) [in Spanish].
10. International Union for Conservation of Nature Species Survival Commission, "Guidelines for reintroductions and other conservation translocations" (2013).
11. F. Sarrazin, J. Lecomte, *Science* **351**, 922 (2016).
12. P. J. Seddon, D. P. Armstrong, R. F. Maloney, *Conserv. Biol.* **21**, 303 (2007).

10.1126/science.abo5970

How to weaken Russian oil and gas strength

Oil and gas exports represent Russia's key geopolitical strength, as well as its major source of hard currency revenues (*1*–*3*). At current prices, these are estimated at around US$1 billion per day (*4*), representing an important lifeline for an economy under heavy financial sanctions in response to Russian President Vladimir Putin's invasion of Ukraine. The European Union (EU) buys 75% of Russian gas exports and 50% of Russian oil exports (*5*). An EU embargo on Russia would substantially undermine Putin's geopolitical and economic positions. However, given a full embargo's potential impacts on the EU, securing the approval of all EU countries is difficult. To limit Russia's oil and gas revenues while keeping up the flows, the EU should instead introduce a tariff on Russian oil and gas imports. The tariff can be adapted to the economic and political dynamics of the conflicts.

A tariff's effect on domestic prices depends on the relative elasticities of supply and demand—i.e., on whether sellers and buyers have relatively better alternatives. The more inelastic the supply (e.g., because Russian exports cannot be diverted) and the more elastic the demand (e.g., because the EU can replace Russian supplies), the more of the tariff will be paid by the supplier (*6*). Russian oil and gas exports to Europe are inelastic in the medium term because infrastructural bottlenecks prevent a substantial redirection to Asia. The EU therefore has a real chance to ensure that tariff revenues are mostly paid by Russia.

To improve its position, the EU needs to increase its demand elasticity. This can be done by incentivizing a reduction of oil and gas demand in Europe and by increasing the use of all available alternative energy resources. By implementing a bold energy strategy, Europe can credibly threaten to cut Russia's oil and gas revenue while minimizing the domestic economic consequences of a tariff.

Ricardo Hausmann[1], Agata Łoskot-Strachota[2], Axel Ockenfels[3], Ulrich Schetter[1], Simone Tagliapietra[4,5], Guntram Wolff[4,6]*, Georg Zachmann[4]
[1]Harvard University, Cambridge, MA 02138, USA. [2]OSW Centre for Eastern Studies, Warsaw, Poland. [3]University of Cologne, Cologne, Germany. [4]Bruegel, 1210 Saint-Josse-ten-Noode, Belgium. [5]Catholic University of the Sacred Heart, Milan, Italy. [6]Free University of Brussels (Université Libre de Bruxelles), Brussels, Belgium.
*Corresponding author.
Email: guntram.wolff@bruegel.org

REFERENCES AND NOTES

1. Ministry of Finance of Russia, Federal Budget (2022); https://minfin.gov.ru/common/upload/library/2022/03/main/fedbud_month_eng.xlsx.
2. Observatory of Economic Complexity, "Russia" (2020); https://oec.world/en/profile/country/rus.
3. Trading Economics, "Russia exports" (2022); https://tradingeconomics.com/russia/exports.
4. International Energy Agency, "Frequently asked questions on energy security" (2022); www.iea.org/articles/frequently-asked-questions-on-energy-security.
5. US Energy Information Administration, "Europe is a key destination for Russia's energy exports" (2022); www.eia.gov/todayinenergy/detail.php?id=51618.
6. P. Mieszkowski, *J. Econ. Lit.* **7**, 1103 (1969).

10.1126/science.abq4436

A global plastic treaty must cap production

In March, the UN Environment Assembly adopted a resolution to combat plastic pollution with a global and legally binding plastics treaty by 2024 (*1*). In his News In Depth story "United Nations to tackle global plastics pollution" (25 February, p. 801), E. Stokstad discusses many of the ambitious provisions that were included, such as a consideration of the whole plastic life cycle and binding targets. However, it is unclear whether the treaty will include a cap on production or cover plastic chemicals. Despite interventions by the industry (*2*) and objections from the United States and other delegations, reducing plastics at the source by curbing production is critical.

The current mass of plastic production is at about 450 million tons annually and set to double by 2045 (*3*). The immense quantity and diversity of both plastics and plastic chemicals, the total weight of which exceeds the overall mass of all land and marine animals (*4*), already poses enormous challenges. Ensuring the safety of every available plastic and chemical is impossible, as their rates of appearance in the environment exceed governments' capacities to assess associated risks and control problems (*5*). Plastic pollutants have altered vital Earth system processes to an extent that exceeds the threshold under which humanity can survive in the future (i.e., the planetary boundary) (*5*). Because legacy plastics in the environment break down into micro- and nanoparticles (*6*), this form of pollution is irretrievable and irreversible (*6*). In addition to the risks for human and environmental health, the whole life cycle of plastic accounts for 4.5% of our current greenhouse gas emissions (*7*) and could consume 10 to 13% of our remaining CO_2 budget by 2050 (*8*). The growing production and inevitable emissions of plastics will exacerbate these problems (*6*).

Failing to address production will lead to more dependence on flawed and insufficient strategies. Some waste management technologies, such as forms of thermal and chemical recycling, cause socioeconomic and environmental harm (*9*). Much of the plastic waste is currently exported from the North to the Global South, which poses a substantial threat to marginalized and vulnerable communities and their environments (*10*). Even when applying all political and technological solutions available today, including substitution, improved recycling, waste management, and circularity, annual plastic emissions to the environment can only be cut by 79% over 20 years; after 2040, 17.3

million tons of plastic waste will still be released to terrestrial and aquatic environments every year (*11*). To fully prevent plastic pollution, the path forward must include a phaseout of virgin plastic production by 2040 (*12*).

Melanie Bergmann[1]*, Bethanie Carney Almroth[2], Susanne M. Brander[3], Tridibesh Dey[4], Dannielle S. Green[5], Sedat Gundogdu[6], Anja Krieger[7], Martin Wagner[8], Tony R. Walker[9]

[1]Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, D-27570 Bremerhaven, Germany. [2]Department of Biological and Environmental Sciences, University of Gothenburg, Gothenburg, Sweden. [3]Department of Fisheries, Wildlife, and Conservation Sciences, Coastal Oregon Marine Experiment Station, Oregon State University, Corvallis, OR 97331, USA. [4]Department of Sociology, Philosophy, and Anthropology, University of Exeter, Exeter EX4 4PY, UK. [5]Applied Ecology Research Group, School of Life Sciences, Anglia Ruskin University, Cambridge CB1 1PT, UK. [6]Faculty of Fisheries, Cukurova University, 01330 Adana, Turkey. [7]Berlin, Germany. [8]Department of Biology, Norwegian University of Science and Technology, Trondheim, Norway. [9]School for Resource and Environmental Studies, Dalhousie University, Halifax, NS B3H 4R2, Canada.
*Corresponding author. melanie.bergmann@awi.de

REFERENCES AND NOTES

1. United Nations Environment Assembly of the United Nations Environment Programme, "End plastic pollution: Towards an international legally binding instrument" (2022); https://wedocs.unep.org/bitstream/handle/20.500.11822/38522/k2200647_-_unep-ea-5-l-23-rev-1_-_advance.pdf?sequence=1&isAllowed=y.
2. J. Geddie, V. Volcovici, J. Brock, M. Dickerson, "U.N. pact may restrict plastic production: Big Oil aims to stop it" (Reuters, 2022).
3. R. Geyer, in *Mare Plasticum—The Plastic Sea: Combatting Plastic Pollution Through Science and Art*, M. Streit-Bianchi, M. Cimadevila, W. Trettnak, Eds. (Springer International Publishing, Cham, 2020), pp. 31–47.
4. E. Elhacham, L. Ben-Uri, J. Grozovski, Y. M. Bar-On, R. Milo, *Nature* **588**, 442 (2020).
5. L. Persson *et al.*, *Environ. Sci. Technol.* **56**, 1510 (2022).
6. M. MacLeod, H. P. H. Arp, M. B. Tekman, A. Jahnke, *Science* **373**, 61 (2021).
7. L. Cabernard, S. Pfister, C. Oberschelp, S. Hellweg, *Nat. Sustain.* **5**, 139 (2022).
8. L. A. Hamilton, S. Feit, "Plastic & Climate: The hidden costs of a plastic planet" (Center for International Environmental Law, Washington, DC, 2019).
9. F. Demaria, S. Schindler, *Antipode* **48**, 293 (2016).
10. C. Wang, L. Zhao, M. K. Lim, W.-Q. Chen, J. W. Sutherland, *Resour. Conserv. Recycl.* **153**, 104591 (2020).
11. W. W. Y. Lau *et al.*, *Science* **369**, 1455 (2020).
12. N. Simon *et al.*, *Science* **373**, 43 (2021).

COMPETING INTERESTS

S.M.B. has served as a cochair and microplastics expert on an advisory panel for the California Ocean Science Trust.

10.1126/science.abq0082

ERRATA

Erratum for the Research Article "Copper induces cell death by targeting lipoylated TCA cycle proteins," by P. Tsvetkov *et al.*, *Science* 376, eabq4855 (2022). Published online 22 April 2022; 10.1126/science.abq4855

Erratum for the Report "A ubiquitous tire rubber–derived chemical induces acute mortality in coho salmon," by Z. Tian *et al.*, *Science* 375, eabo5785 (2022). Published online 18 February 2022; 10.1126/science.abo5785

Where Science Gets Social.

AAAS.ORG/COMMUNITY

AAAS' Member Community is a one-stop destination for scientists and STEM enthusiasts alike. It's "Where Science Gets Social": a community where facts matter, ideas are big and there's always a reason to come hang out, share, discuss and explore.

Member COMMUNITY AAAS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE

RESEARCH

IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

MARINE BIOLOGY

Local adaptation in coral reefs

The Clerke Reef atoll at the Rowley Shoals in Western Australia provides a model habitat with which to study how spatially varying selection drives ecological divergence. Studying the broadcast spawning coral *Acropora tenuis*, Thomas *et al.* examined individuals from lagoon and reef-flat locations through genome and transcriptome sequencing combined with common garden heat-stress experiments. Corals in the lagoon were shown to be more heat tolerant, and the gene expression baseline also differed between the two sites. Understanding existing genetic variation across habitats could help to protect and manage reefs in the upcoming decades. —MM *Sci. Adv.* 10.1126/sciadv.abl9185 (2022).

Spatial variation in lagoon habitats at the Clerke Reef atoll has given rise to corals with increased tolerance to heat.

PHOTO: SUZANNE LONG/ALAMY STOCK PHOTO

NEUROSCIENCE

The spinal circuit for pleasant touch

Pleasant touch (e.g., cuddling, caressing, and patting) encodes positive hedonic information that facilitates emotional bonding, affiliative behavior, and the well-being of social animals. Despite its profound importance, how pleasant touch information is encoded and transmitted from sensory neurons to the spinal cord remains unknown. Liu *et al*. identified interneurons in laminae II of the spinal cord dorsal horn that express prokineticin receptor 2, as well as sensory neurons that express the ligand prokineticin 2 that are involved in the encoding and transmitting of pleasant touch. Genetic ablation of these neurons selectively abolished the place preference that mice developed in a pleasant touch–conditioned place preference test. However, the sensation of pain and itch remained unaffected in the mutant animals. —PRS

Science, abn2479, this issue p. 483

TUMOR IMMUNOLOGY

A CAR for killing, not toxicity

Chimeric antigen receptor T (CAR-T) cell therapies are approved by the US Food and Drug Administration for treating blood cancers but are often limited by toxicities and difficult and expensive production. Developing more-streamlined CAR-T cell therapies is crucial. Vasic *et al.* used mouse and human double-negative (DN) T cells as a CAR-T cell therapy, which showed robust efficacy in mouse models of blood and lung cancers. They found that the DN CAR-T cells were just as effective as conventional CAR-T cells but induced no toxicity. The DN CAR-T cells could be produced from mixed donors, were not rejected from hosts, and could be frozen for long periods of time, suggesting that they could be an attractive "off-the-shelf" CAR-T cell therapy option. —DAE

Sci. Immunol. **7**, eabl3642 (2022).

HYDROGELS

Converting forces into ionic signals

Piezoelectric materials can convert mechanical forces into electrical signals and are used as the active material in many pressure sensors. However, biological systems tend to be based on the movement of ions rather than electrons. Dobashi *et al*. looked at the piezoionic effect of several kinds of hydrogels. Hydrogels are designed so that the anions and cations

have a different mobility; thus, when the material is squeezed, it causes an ionic gradient that generates voltage. The authors demonstrate several potential applications, including a piezoionic skin and peripheral nerve stimulation, to demonstrate the possibility of self-powered piezoionic neuromodulation. —MSL

Science, aaw1974, this issue p. 502

CANCER

Self-inflicted damage protects tumors

Genotoxic treatments, such as radiation and some chemotherapy drugs, are a mainstay of cancer therapy, but they often fail to fully destroy tumor cells. Normal cells can protect themselves from genotoxic insults by activating the G_1 cell cycle checkpoint, but this checkpoint is often dysfunctional in tumors. By contrast, Larsen *et al*. discovered that tumor cells can activate a nuclease that causes limited induction of DNA breaks at specific sites, which is coordinated with the process of DNA break repair. These self-inflicted DNA breaks trigger the G_2 cell cycle checkpoint, preventing tumor cells from cycling and protecting them from death due to treatment-induced DNA damage. —YN

Science, abi6378, this issue p. 476

DYNAMIC GENOME

Fleeting chromatin loops

The genome is organized into three-dimensional (3D) domains that are widely thought to be stable, fully looped structures, although this organization has not been directly observed in living cells. Gabriele *et al*. report the direct visualization of chromatin looping in living cells and used Bayesian inference to quantify looping dynamics. Loops were found to be both rare and short-lived, overturning static models of looping. Instead of being fully looped, 3D genome domains existed overwhelmingly in partially folded configurations. This more dynamic view of 3D genome domains may ultimately allow a deeper understanding of why the disruption of some domains and loops causes dysregulation of gene expression in disease. —DJ

Science, abn6583, this issue p. 496

GRAVITATIONAL WAVES

A gamma-ray pulsar timing array

After galaxies merge, the supermassive black holes (SMBHs) at their centers are expected to form binaries that emit gravitational waves at nanohertz frequencies. Numerous SMBH binaries throughout the Universe should combine to produce a gravitational wave background. Existing searches for this signal use radio observations of pulsars as sensitive clocks and look for small shifts in the pulse timings. The Fermi-LAT Collaboration implemented a pulsar timing array using gamma rays and achieved a sensitivity close to that of the radio approaches. The results set an independent upper limit on the gravitational wave background, which is subject to different noise sources. —KTS

Science, abm3231, this issue p. 521

ORGANIC CHEMISTRY

Carbon excision

Quinolines and indoles are both very common core motifs in drug molecules. Because they differ by just a single carbon in their ring frameworks, it would be useful to interconvert them during structure-activity relationship studies. However, aromatic stabilization in these compounds makes that process difficult. Woo *et al*. report that narrow-wavelength irradiation of quinoline *N*-oxides in the near-ultraviolet range followed by acid treatment cleanly excises a ring carbon to produce an indole, avoiding secondary photoproducts that were previously observed using broadband light. —JSY

Science, abo4282, this issue p. 527

IN OTHER JOURNALS

Edited by **Caroline Ash** *and* **Jesse Smith**

URBAN ECOLOGY

Embracing the nightlife

Some animals expertly adapt to urban environments, often by changing their behavior. Gallo *et al.* investigated whether mammals in cities change the timing of their activities to earlier or later in the day. Using images from camera traps in 10 US cities, the authors analyzed how mammal activity changed across gradients in human population density, impervious surface cover, temperature, and vegetation. Coyotes were more nocturnal in more populated areas. Common omnivore and herbivore species changed their activity with urbanization, although some became more nocturnal and some less. Coyotes and red foxes avoided hot temperatures by increasing activity at dusk. Such temporal shifts can allow animals to exploit resources, avoid conflict, and navigate urban environments. —BEL *eLife* **11**, e74756 (2022).

Urbanized areas attract certain wild species, many of which have changed their behavior to exploit new opportunities, like this raccoon foraging in a garbage bin at night.

PHOTO: MIRCEAX/ISTOCK.COM

SCIENCE EDUCATION

Loading the dice for active learning

The benefits of active learning have been well established, but traditional lecture-orientated approaches are still the norm at many institutions. Yik *et al.* implemented large-scale data collection (from 2382 instructors) to empirically test the relative association of 17 malleable factors that influence the uptake of active learning. Using regression analysis and multilevel modeling, the team found that contextual factors (e.g., classroom setup) and personal contexts (e.g., participation in scholarship of teaching and learning activities) were associated with a decrease in the percentage of time spent lecturing. Broad recommendations for sustaining active learning strategies are discussed, including constructing classroom spaces that support active learning, incentives for professional development, and meaningful recognition of scholarship of teaching and learning work accomplished by instructors. —MMc

Int. J. STEM Educ. **9**, 15 (2022).

NONCOVALENT POLYMERS

Photocontrol fluttering controls polymerization

Most methods for pausing and resuming polymerization reactions act at the growing chain. Aoki *et al.* show how light-induced conformer changes of a monomer switch the kinetics of a supramolecular polymerization. A cyclooctatetraene derivative, formed from fused thiophene molecules that bear hydrogen-bonding side chains, has a chiral nonplanar central ring that can racemize through a fluttering motion. This molecule polymerizes in the dark as alternating *R* and *S* monomers but not during prolonged ultraviolet irradiation.

In the dark, the fluttering motion is much slower than chain elongation, but under irradiation, fluttering is 11 times faster. This motion inhibits both nucleation and polymerization until the solution is returned to the dark state and provides a photoswitch for polymerization. —PDS

J. Am. Chem. Soc. 10.1021/jacs.2c02176 (2022).

CANCER GENOMICS

Mind the genome gaps

The Cancer Genome Atlas (TCGA) is a major resource in cancer genomics, encompassing sequencing data for 33 different human cancer types. Unfortunately, ethnic minority patients are underrepresented in TCGA, even though they often have a higher cancer risk. Wickland *et al*. examined TCGA data for self-reported Black patients with seven tumor types and discovered that for six of these cancers, the tumor genomes of these patients were sequenced at a lower depth than tumor genomes of those with European ancestry. In some cases, this happened because the genomes of most patients with African ancestry were sequenced in later batches. Additionally, they were more likely to have insufficient coverage in parts of the exome, likely because the human reference genome is biased toward European ancestry. These findings highlight important gaps to address in future sequencing efforts. —YN

J. Natl. Cancer Inst. 10.1093/jnci/djac054 (2022).

NEUROSCIENCE

Unwrapping TDP-43

Oligodendrocytes continually renew and mature to generate myelin sheaths that enwrap and insulate axons. TDP-43 is a protein known to play a role in neurodegenerative disease. Heo *et al.* studied TDP-43 loss of function in the mouse brain. Selective deletion of TDP-43 from oligodendrocyte precursor cells resulted in their rapid death. Deletion of TDP-43 at different stages of oligodendrocyte maturation led to different outcomes. Early deletion caused degeneration of mature oligodendrocytes, seizures, and premature lethality, whereas late deletion spared oligodendrocytes and allowed seizure-free survival. Both manipulations led to thinner and fewer myelin sheaths and to inappropriate wrapping of neuronal cell bodies and blood vessels. Conditional removal of TDP-43 from oligodendrocytes in adults caused hindlimb weakness. Loss of TDP-43 led to the missplicing of RNAs that encode key genes involved in oligodendrocyte morphogenesis and myelination and may underlie abnormal neuronal activity and neurodegeneration in diverse neurological diseases. —SMH

eLife **11**, e75230 (2022).

POLARITONIC CHEMISTRY

Proton transfer inside an optical cavity

There is considerable interest in developing approaches to controlling proton transfer reactions, which play a key role in a wide range of practically important energy conversion processes. Using quantum electrodynamics ab initio methods, Pavošević *et al.* studied the effect of strong light-matter interactions, which can be created by an optical cavity, on intramolecular proton transfer in malonaldehyde and aminopropenal. The authors showed that the optical cavity can increase or decrease the rate of proton transfer in these two prototypical systems, depending on the cavity-mode polarization. The present work demonstrates that the effect of an optical cavity could serve as a practical route to controlling proton transfer reactions and should trigger further experimental developments in this direction. —YS

J. Am. Chem. Soc. **144**, 4995 (2022).

INNATE IMMUNITY

Don't hold your breath!

Breathing does more than just ensure the oxygenation of an animal's tissues. With its dynamic and regular mechanical forces, breathing has been implicated in all sorts of processes, including lung development and regulation of immune responses. Bai *et al.* used a human lung alveolus chip microfluidic device that recapitulates human airway physiology to model H3N2 influenza infection. They found that cyclical respiratory motion inhibits viral infection by activating the mechanosensitive ion channel TRPV4, which then induces lung epithelium and endothelium to produce S100 alarmin proteins. These mediators bind the pattern recognition receptor RAGE, which triggers several innate immune response programs that help combat virus. Intriguingly, a TRPV4 inhibitor was able to reduce both inflammation and viral loads, suggesting that this may be a useful approach for the treatment of viral pneumonia. —STS

Nat. Commun. **13**, 1928 (2022).

ALSO IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

PAIN

Trk'ing mutations for pain relief

The nerve growth factor receptor tropomyosin receptor kinase A (TrkA) is an attractive target for alleviating pain, but it also mediates other signaling pathways that are critical to neuronal function and survival. To find targets specifically within its pain-mediating pathways, Moraes *et al*. examined the structural consequences of TrkA mutations that cause a pain-insensitivity disorder. They found mutations that disrupted the interaction of TrkA with one of its downstream effectors, phospholipase Cγ. Treating mice with a peptide designed to similarly disrupt the interaction reduced sensitivity to painful inflammation. —LKF

Sci. Signal. **15**, eabm6046 (2022).

DOG GENOMICS

What is your dog like?

Modern domestic dog breeds are only ~160 years old and are the result of selection for specific cosmetic traits. To investigate how genetics align with breed characteristics, Morrill *et al*. sequenced the DNA of more than 2000 purebred and mixed-breed dogs. These data, coupled with owner surveys, were used to map genes associated with behavioral and physical traits. Although many physical traits were associated with breeds, behavior was much more variable among individual dogs. In general, physical trait heritability was a greater predictor of breed but was not necessarily a predictor of breed ancestry in mutts. Among behavioral traits, biddability—how well dogs respond to human direction—was the most heritable by breed but varied significantly among individual dogs. Thus, dog breed is generally a poor predictor of individual behavior and should not be used to inform decisions relating to selection of a pet dog. —LMZ

Science, abk0639, this issue p. 475

INFECTIOUS DISEASE

Another pandemic threat?

Over the past year, a highly pathogenic avian influenza virus (HPAIv) lineage has caused widespread outbreaks and deaths in poultry and wild birds, as well as some infections in humans. Although there is no evidence of human-to-human transmission, there is concern that the prevalence of these viruses could result in adaptation that allows transmission between humans. In addition, the unprecedented spread and frequency in Europe, Asia, and North America has had major impacts on the poultry industry and is a threat to wild birds, especially those that are already endangered. In a Perspective, Wille and Barr discuss the factors that have led to outbreaks of this HPAIv lineage, the consequences, and what can be done. —GKA

Science, abo1232, this issue p. 459

CANCER

Building up to melanoma

Numerous cancer-causing mutations have been identified over the years, but they do not occur in isolation, and it is difficult to disentangle the effects of individual genetic changes in a tumor that contains hundreds of mutations. To gain insight into melanoma biology and facilitate further study of this cancer, Hodis *et al*. began with healthy human melanocytes and sequentially introduced melanoma-associated mutations using gene editing. The edited cells were then grown in mouse models, simulating human melanomas with different mutation patterns. The authors examined the genomics, histology, and biological behavior of the resulting tumors and compared them with naturally occurring human tumors, demonstrating the relevance of their approach. —YN

Science, abi8175, this issue p. 474

PATHOGEN EVOLUTION

Immunity-induced evolution

Population immunity is thought to drive pathogen evolutionary dynamics, but this process can be hard to measure. Considering the switch from whole-cell vaccines to acellular vaccines (ACVs) as a natural experiment, Lefrancq *et al*. analyzed the diversity of *Bordetella pertussis* genome isolates from five continents to infer local changes in strain fitness after ACV introduction. Their analysis provides evidence that ACVs have led to the increased fitness of strains that lack the virulence factor pertactin and may have implications for understanding asymptomatic *B. pertussis* infection. —CAC

Sci. Transl. Med. **14**, eabn3253 (2022).

BATTERIES

Observations of cathode evolution

The loss of capacity in a rechargeable battery can be due to changes in the electrode structure that occur with cycling. Li *et al*. used hard x-ray holotomography to visualize the structure of a nickel-rich $LiNi_{0.8}Mn_{0.1}Co_{0.1}O_2$ composite cathode (see the Perspective by Xiao). They were able to track the behavior of thousands of individual particles with time and thus determine the relationship between structure and performance as well as the deterioration of the cathode at a size scale that is not generally accessible. They found that damage during cycling is driven not only by each particle but also by its surrounding neighbors, although the contributions shift over time. This work suggests ways to better design electrodes to maximize their performance. —MSL

Science, abm8962, this issue p. 517; see also abo7670, p. 455

ZEOLITES

Benzene stretches zeolite channels

The pores of zeolites allow for selective adsorption of molecules based on their size and shape. The pore sizes that are calculated based on their structure are smaller than what would be expected from their absorption and chemical reaction properties, which implies that the frameworks must have inherent flexibility. Xiong *et al*. used environmental transmission electron microscopy to image the straight channels of ZSM-5 zeolite with adsorbed benzene (see the Perspective by Willhammar and Zou). The pores stretched along the longest direction of the confined benzene molecules by up to 15%. This large change was compensated for by adjacent channels so that the overall deformation of the original unit cells was less than 0.5%. —PDS

Science, abn7667, this issue p. 491; see also abo5434, p. 457

EXTINCTION

Rising temperatures, rising risks

Climate change brings with it the increasing risk of extinction across species and systems. Marine species face particular risks related to water warming and oxygen depletion. Penn and Deutsch looked at extinction risk for marine species across climate warming and as related to ecophysiological limits (see the Perspective by Pinsky and Fredston). They found that under business-as-usual global temperature increases, marine systems are likely to experience mass extinctions on par with past great extinctions based on ecophysiological limits alone. Drastically reducing global emissions, however, offers substantial protection, which emphasizes a need for rapid action to prevent possibly

catastrophic marine extinctions. —SNV

Science, abe9039, this issue p. 524; see also abo4259, p. 452

NEUROETHOLOGY

Bees want, too

In mammals, the subconscious "wanting" of something, such as food, is facilitated by a complex neurobiological process that involves dopaminergic signaling. In a eusocial animal, such as the honey bee, individuals supply food not just for themselves but also for their colony. Huang *et al*. looked at the neurobiological basis of wanting in honey bees (*Apis mellifera*) and found that it is also controlled by a dopamine-based signaling process, suggesting a shared mechanism across millions of years of divergence (see the Perspective by Garcia and Dyer). Furthermore, they found that wanting was stimulated in bees based on their own individual desire to forage as well as observation of the bee waggle dance, suggesting the existence of a colony-level motivation mechanism. —SNV

Science, abn9920, this issue p. 508; see also abp8609, p. 456

EPIDEMIOLOGY

The madness of dogs

Rabies is a deadly zoonotic disease that causes tens of thousands of deaths every year, mainly among African and Asian children. It is usually transmitted to humans by dogs, and once symptoms appear, it is inevitably fatal. Despite control efforts, rabies continues to circulate at very low prevalence in Serengeti district, Tanzania, where Mancy *et al*. have been tracking cases in several species for more than a decade to understand the dynamics of the virus. Highly resolved genetic data allowed the authors to trace transmission networks, from which individual dog behavior emerges as a key factor (see the Perspective by Antolin). Some infected dogs may travel long distances and introduce new lineages into neighboring communities, whereas others may simply bite other animals. Because dogs are mobile, culling fails, and the only resort is fully comprehensive dog vaccination. —CA

Science, abn0713, this issue p. 512; see also abo7428, p. 453

ORGANIC CHEMISTRY

Additive improvements to nickel catalysis

It often takes decades of incremental optimization to apply chemical reactions beyond the small range of substrates studied at the discovery stage. Prieto Kullmer *et al*. sought to accelerate that optimization process by screening a large, diverse group of additives to a cooperative nickel-photoredox catalyst system. The screen revealed that phthalimides substantially expand the functional compatibility of the nickel catalyst and thus the substrate scope. The phthalimide appears to stabilize oxidative addition complexes as well as break up deactivated catalyst aggregates. —JSY

Science, abn1885, this issue p. 532

RESEARCH ARTICLE SUMMARY

CANCER

Stepwise-edited, human melanoma models reveal mutations' effect on tumor and microenvironment

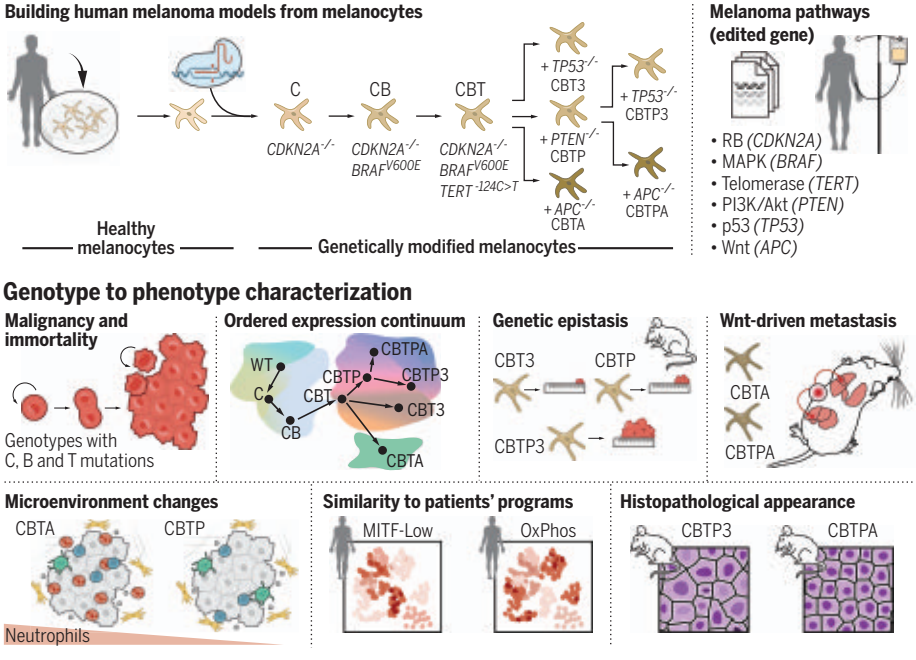
Eran Hodis*, Elena Torlai Triglia, John Y. H. Kwon, Tommaso Biancalani, Labib R. Zakka, Saurabh Parkar, Jan-Christian Hütter, Lorenzo Buffoni, Toni M. Delorey, Devan Phillips, Danielle Dionne, Lan T. Nguyen, Denis Schapiro, Zoltan Maliga, Connor A. Jacobson, Ayal Hendel, Orit Rozenblatt-Rosen, Martin C. Mihm Jr., Levi A. Garraway, Aviv Regev*

INTRODUCTION: Cancer develops from cells that become malignant because of mutations in multiple genes—often accumulated over long time periods—that produce a phenotypic diversity across patient tumors. Specific genetic alterations in particular cancer types have been linked to prognosis, to response or resistance to therapies (especially targeted therapeutics), and to a tumor's propensity to acquire further mutations, among other phenotypes. However, genotype-phenotype connections are challenging to infer in patients, as any two tumors differ genetically in too many ways to isolate the effect of one or several mutations. The ability to systematically connect cancer-associated mutations or combinations thereof with their phenotypic consequences would advance our understanding of the mechanisms of cancer pathogenesis and genetically linked disease features.

RATIONALE: We reasoned that genome editing and the fitness advantage of cancer-associated mutations could be leveraged to generate human cellular models of tumor development. Such genome-edited models would replicate the precise genetics, lineage relationships, and stepwise progression of cancer and allow us to establish genotype-to-phenotype links in a controlled experimental design. While similar models have been realized for tumors originating from self-renewing stem cells as cells of origin, specifically in colorectal cancer, no comparable models exist for tumor types that arise from nonstem differentiated cells. We present an approach that starts from the nonstem cell of origin of melanoma, the healthy human melanocyte; this approach then generates a series of cells with precise genome editing of mutations in key cancer genes, thus expanding the horizon of possible cellular models of cancer development.

RESULTS: We generated a progressive series of genome-edited human models of melanoma. We started from healthy human melanocytes and introduced, in a stepwise fashion, mutations in up to five genes spanning six pathways commonly dysregulated in melanoma: *CDKN2A* (part of the RB pathway), *BRAF* (MAPK), *TERT* (telomerase), *PTEN* (PI3K/AKT), *TP53* (p53), and *APC* (Wnt), for a total of nine genetically distinct cellular models. We characterized these models during growth in vitro and after intradermal injection through mouse xenografts, using physiological assessment, histopathology, and single-cell RNA sequencing (scRNA-Seq), leveraging computational methods and machine learning algorithms. Through these models, we connected melanocyte genotypes to phenotypes such as gene expression programs, replicative immortality, malignancy, rapid tumor growth, tumor pigmentation, metastasis, and histopathological features. In vitro, consecutive mutations produced an ordered progression through expression space. In vivo, mutations in malignant cells also affected the cell-type composition and expression states of tumor-infiltrating microenvironment cells. Our melanoma models shared genotype-associated expression programs with patient melanomas and partially recapitulated patient melanoma genotype-associated histopathological features.

CONCLUSION: The genotype-phenotype connections we identify highlight how the impact of mutations often depends on genetic context. Such genetic epistasis makes understanding the phenotypic consequences of the mutational landscape of human cancers a combinatorial problem whose study requires modeling strategies that can scale to multiple mutations, such as the one presented in this study. Genome-edited human models of cancer enable the identification of causal relationships between defined sets of genetic alterations and disease-relevant phenotypes, furthering understanding of how cancer mutations help give rise to the diverse and varied phenotypes of human malignancy. ■



Genome-engineered human cell models connect melanoma genotypes to phenotypes. Sequential, precise gene editing of human melanocytes produced a series of melanoma models. The editing approach we developed—without selection markers or single-cell cloning—is applicable to other cell types, expanding the toolkit of available cancer models. Through in-depth phenotypic characterization of in vitro cells and in vivo tumors from xenografts in immunodeficient mice, multimutant melanoma genotypes were causally linked to specific phenotypes.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: eran_hodis@dfci.harvard.edu (E.H.); aviv.regev.sc@gmail.com (A.R.)
Cite this article as E. Hodis *et al.*, *Science* **376**, eabi8175 (2022). DOI: 10.1126/science.abi8175

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abi8175

RESEARCH ARTICLE

CANCER

Stepwise-edited, human melanoma models reveal mutations' effect on tumor and microenvironment

Eran Hodis[1,2,3]*†‡§, Elena Torlai Triglia[1]†, John Y. H. Kwon[1,2], Tommaso Biancalani[1]#, Labib R. Zakka[4,5], Saurabh Parkar[1], Jan-Christian Hütter[1]#, Lorenzo Buffoni[1], Toni M. Delorey[1], Devan Phillips[1]#, Danielle Dionne[1], Lan T. Nguyen[1], Denis Schapiro[1,6]¶, Zoltan Maliga[6], Connor A. Jacobson[6], Ayal Hendel[7], Orit Rozenblatt-Rosen[1]#, Martin C. Mihm Jr.[4,5], Levi A. Garraway[1,2]#, Aviv Regev[1,8]*#

Establishing causal relationships between genetic alterations of human cancers and specific phenotypes of malignancy remains a challenge. We sequentially introduced mutations into healthy human melanocytes in up to five genes spanning six commonly disrupted melanoma pathways, forming nine genetically distinct cellular models of melanoma. We connected mutant melanocyte genotypes to malignant cell expression programs in vitro and in vivo, replicative immortality, malignancy, rapid tumor growth, pigmentation, metastasis, and histopathology. Mutations in malignant cells also affected tumor microenvironment composition and cell states. Our melanoma models shared genotype-associated expression programs with patient melanomas, and a deep learning model showed that these models partially recapitulated genotype-associated histopathological features as well. Thus, a progressive series of genome-edited human cancer models can causally connect genotypes carrying multiple mutations to phenotype.

Healthy human cells become cancerous through the acquisition of genetic mutations. Efforts to sequence the genomes of human cancer cells have illuminated the daunting array of mutation combinations that can cause life-threatening malignancies, even when they arise from the same cell type of origin (*1*, *2*). Additionally, a great phenotypic diversity, both within and between patients, is caused in part by the somatic mutations observed in these intricate genetic maps of cancer (*1*, *2*). For example, certain genetic differences may explain why some cancer cells are more prone to metastasize (*3*, *4*), some are less susceptible to immune attack (*5*), and others have genomes that are more likely to accumulate chromosomal alterations (*6*). Thus, linking maps of cancer mutations to disease-relevant phenotypic consequences advances our understanding of cancer biology and may inform the design of genetically-targeted therapies.

However, genotype-to-phenotype connections are not easily revealed by comparing individual human cancers, as any two patient tumors or cell lines typically differ genetically in too many ways to distinguish the effect of a single mutation or a particular combination of mutations (*7*–*9*). Furthermore, early stages in cancer development are rarely represented in patient-derived tumors and cell lines (*10*, *11*). One solution—made possible by recent advances in mammalian genome editing (*12*, *13*)—is to use human cell models to replicate, in an isogenic fashion, precise multimutant genetics, cell lineages, and stepwise progression of cancer. Such human models have been realized for colorectal cancer through the use of stem cells and timed withdrawal or addition of mutation-matched growth factors or chemicals (*14*, *15*). However, there remains a need for an approach that does not depend on foreknowledge of selective pressures nor on single-cell cloning, thus enabling generalized multimutation modeling of nonstem cells.

Melanoma provides a prime case in point for multimutation cancer modeling. Its genetic landscape is both well-charted and complex as a result of sunlight-induced DNA damage (*8*, *16*–*26*). Despite their complexity, nearly all melanomas arising in hair-bearing skin have genetic alterations in the retinoblastoma tumor suppressor (RB) pathway, the mitogen-activated protein kinase (MAPK) pathway, and in telomerase regulation (*20*, *27*). These three molecular pathways are most commonly dysregulated by inactivating mutations or loss of *CDKN2A*, an activating mutation in *BRAF*, and one of two point mutations in the *TERT* promoter, respectively. Melanoma progression is further associated with mutations in many different pathways, including the phosphatidylinositol-3-kinase (PI3K)/Akt pathway (mutated in ~25% of thick melanomas), the p53 pathway (~25% of thick melanomas), and the Wnt pathway (*APC* alterations in ~2 to 7% of melanomas and *CTNNB1* mutations in ~5% of melanomas) (*16*, *27*). Moreover, the melanoma cell of origin—the melanocyte, a pigment-producing skin cell—is known and readily grown in culture. Additionally, primary human melanocytes are amenable to genome editing, and a single melanoma-associated mutation can undergo positive selection in standard melanocyte culture conditions (*28*). Recently, pioneering work has taken the first steps toward genome-engineered melanomas by introducing inactivating *CDKN2A* and activating [Val^{600}→Glu (V600E)] *BRAF* mutations into human melanocytes (*28*). Taken together, these features make melanoma a compelling case study for the development of multimutation cell models. We took an engineering approach that leverages advances in genome editing and the fitness benefit conferred by cancer-associated mutations to generate a collection of multimutation primary cell models, which we characterized molecularly and phenotypically.

Mutation fitness advantage enables multistep genome editing

We developed a strategy to sequentially introduce different series of cancer-associated mutations into healthy, differentiated human cells in culture (Fig. 1A). Exploiting the fitness advantage of cancer-associated mutations (*29*), we repeatedly selected for desired mutations at the endogenous gene loci by waiting for the faster-growing mutant cell population to outcompete nonmutant cells over time in culture, without selection markers or single-cell cloning. To introduce each precise mutation, we delivered the necessary genome editing machinery to cells in vitro by electroporation of Cas9 ribonucleoprotein (RNP) complex and then monitored the mutant allele frequency in the cell population over time under standard cell culture conditions (*30*). Once the mutant alleles surpassed the nonmutant alleles in frequency—indicating that the mutant cells had outcompeted the nonmutant cells and become the predominant population—we introduced the next mutation. We repeated the entire process multiple times to sequentially introduce up to five cancer-associated mutations per cell model.

We created a melanocyte genome-editing tree guided by both human melanoma genetics

[1]Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA. [2]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA 02115, USA. [3]Harvard-MIT Division of Health Sciences and Technology, Harvard Medical School, Boston, MA 02115, USA. [4]Department of Dermatology, Brigham and Women's Hospital, Boston, MA 02115, USA. [5]Harvard Medical School, Boston, MA 02115, USA. [6]Laboratory of Systems Pharmacology, Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA. [7]The Mina and Everard Goodman Faculty of Life Sciences and Advanced Materials and Nanotechnology Institute, Bar-Ilan University, Ramat-Gan 52900, Israel. [8]Koch Institute for Integrative Cancer Research, Department of Biology, MIT, Cambridge, MA 02139, USA.
***Corresponding author. Email: eran_hodis@dfci.harvard.edu (E.H.); aviv.regev.sc@gmail.com (A.R.)**
†These authors contributed equally to this work.
‡Present address: Department of Medicine, Brigham and Women's Hospital, Boston, MA 02115, USA.
§Present address: Harvard Medical School, Boston, MA 02115, USA.
¶Present address: Institute for Computational Biomedicine and Institute of Pathology, Faculty of Medicine, Heidelberg University Hospital and Heidelberg University, 69120 Heidelberg, Germany.
#Present address: Genentech, 1 DNA Way, South San Francisco, CA 94080, USA.

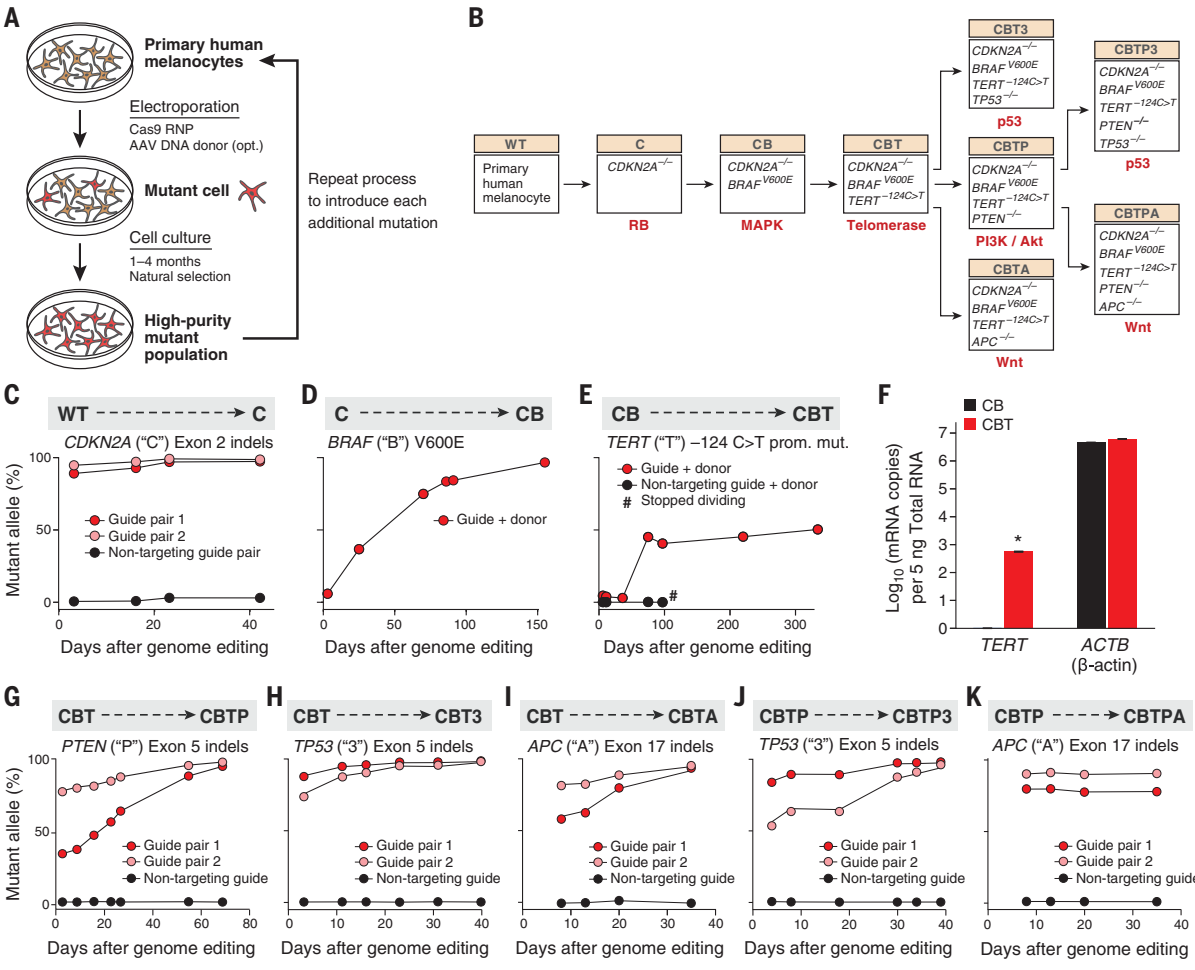


Fig. 1. Fitness advantage of cancer-driving mutations enables the creation of a progressive series of genome-edited human cancer models. (**A**) Experimental approach for introducing sequential melanoma mutations into the genomes of primary human melanocytes with CRISPR-Cas9. RNP, ribonucleoprotein; AAV, adeno-associated virus. (**B**) Editing tree showing the nine isogenic models of melanoma generated (boxes), the perturbed genes in each model (white boxes), the genotype abbreviation (beige boxes), and the molecular pathway dysregulated by the most recent genome edit (red text). (**C** to **E**) Sequential introduction of first three mutations by CRISPR-Cas9 genome editing of wild-type (WT) melanocytes. (C) First mutation: *CDKN2A* ("C"). (D) Second mutation: *BRAF* ("B"). (E) Third mutation: *TERT* ("T"). *TERT* editing confers replicative immortality to CB melanocytes. Allele frequencies of each engineered mutation (*y* axis) shown over time (*x* axis). #, measurement of allele frequency discontinued because of cell senescence. (**F**) Addition of the -124C>T *TERT* promoter mutation activates *TERT* expression. Mean of log 10 number of *TERT* and β-actin (*ACTB*) transcripts (*y* axis) measured by qPCR in CB (black) and CBT (red) cells. Error bars, SD. $n = 3$. *, $P < 0.01$, one-tailed, one-sample Student's *t* test. (**G** to **I**) Introduction of fourth mutation into CBT melanocytes. (G) Allele frequencies of knockout of *PTEN* ("P"), (H) knockout of *TP53* ("3"), and (I) knockout of *APC* ("A"). (**J** and **K**) Introduction of fifth mutations into CBTP melanocytes (J) Allele frequency of knockout of *PTEN* and (K) knockout of *TP53*. Allele frequencies (*y* axis) shown over time (*x* axis), as assessed by indels in the respective loci in genomic DNA.

and technical complexity (Fig. 1B), by sequentially introducing mutations into the RB pathway, then the MAPK pathway, and then the telomerase regulation pathway through precise editing of healthy human melanocytes. We knocked out the *CDKN2A* ("C", RB pathway) by electroporating a genome-editing Cas9 RNP targeting the *CDKN2A* locus (exon 2, shared by both of its protein products, p16 and p14) (*30*). Small insertions and deletions (indels) in the gene underwent positive selection in culture, reaching 90 to 95% mutated allele frequency by day 3 and 98 to 99% by day 42 [Fig. 1C; mutated allele frequency quantified as percent of alleles with an indel; the two predominant alleles were a 79– and 53–base pair (bp) deletions between the guides of pairs 1 and 2 at 88 and 75%, respectively]. We next introduced the BRAF V600E mutation ("B", MAPK pathway) into C-edited melanocytes by codelivering Cas9 RNP targeting *BRAF* exon 15 and a homologous DNA donor encoding the V600E mutation (*30*). Recombinant adeno-associated virus (AAV) was used to deliver the DNA donor to overcome the low editing efficiency of the single-stranded oligodeoxynucleotide donors (<0.25% at day 6; table S1) (*31*). The *BRAF* V600E allele frequency increased from 6% at day 3 to 97% at day 155 in culture (Fig. 1D, reflecting a homozygous population and likely indicating that, in this context, two V600E alleles provide a greater fitness advantage than one allele). Finally, we introduced the *TERT* -124C>T promoter mutation ("T", telomerase regulation) into "CB" melanocytes, codelivering a Cas9 RNP targeting *TERT* exon 1 and a homologous DNA donor encoding *TERT* -124C>T. The frequency of the -124C>T *TERT* mutated allele increased from 3 to 5% in the first 30 days in culture to 45% by day 75 and stayed at 41 to 50% for more than 300 days of continuous culture (Fig. 1E; ascertained as predominantly heterozygous with a small subpopulation of homozygous cells) (*30*). Although the more common mutation

trajectory in human melanoma may be B → T → C (*18*, *27*), we opted to engineer the path C → B → T to defer the more difficult, precise genome edits to a stage when the cells would be more cancerlike and therefore more amenable to editing (*32*) (the B → T → C order was not attempted, though previous work has demonstrated the feasibility of engineering B as the first mutation) (*28*). Engineering the *TERT* promoter mutation was the most technically difficult of the three mutations; it required testing 40 different Cas9 guide sequences to identify a potent reagent for making double stranded breaks near the *TERT* promoter locus (table S2). This difficulty was possibly due to the high G:C content or closed chromatin state at this locus (*33*). We termed the resulting triple-mutation cells CBT melanocytes.

This first progressive series of mutant melanocyte models addressed whether *TERT* promoter mutations turn on *TERT* expression and confer replicative immortality in the appropriate genetic and cellular context. Indeed, the CBT melanocytes showed *TERT* expression by quantitive polymerase chain reaction (qPCR), whereas the CB melanocytes exhibited none (Fig. 1F). Furthermore, the CBT melanocytes grew indefinitely in culture (>1.5 years), whereas the CB cells exhibited morphological signs of senescence ("fried egg" appearance; fig. S1) and stopped dividing by day 100 (Fig. 1E, black curve and hash marks), by which point the cells had been in continuous culture for ~6 months since the original thaw of the wild-type (WT) parental melanocytes. We observed comparable effects on *TERT* expression and replicative immortality with the other common *TERT* promoter mutation, -146C>T (*21*, *22*) (fig. S2). Thus, either -124C>T or -146C>T *TERT* promoter mutation is sufficient to activate *TERT* expression and confer replicative immortality upon human melanocytes in the CB genetic context.

Because melanoma progression is associated with mutations in many different pathways—including the PI3K/Akt, p53, and Wnt pathways (*16*, *27*, *34*)—we explored the effect of subsequent loss-of-function mutations in *PTEN* ("P"), *TP53* ("3"), or *APC* ("A"), respectively (Fig. 1B). Indels in each of the fourth targeted genes underwent positive selection in culture, reaching 94 to 99% mutant allele frequency by 70 days at most, separately yielding CBTP, CBT3, and CBTA melanocytes (Fig. 1, G to I).

Finally, to explore combinations of melanoma progression mutations, we extended the P branch of the editing tree by introducing either the 3 or A mutation (Fig. 1B). In producing CBTP3 melanocytes, indels in *TP53* increased over time in culture to an allele frequency of 96 to 98%, (Fig. 1J). In producing CBTPA melanocytes, indels in *APC* stayed at a stable allele fraction of ~75 to 85% (Fig. 1K) only to later increase to >99% when grown in vivo in xenograft studies (six of six examined tumors, see below). Upon introduction of each mutation throughout the editing tree, we observed the expected functional effect on the relevant molecular pathway, confirming the activity of the genetic mutations (fig. S3).

Our genome-edited tree of human melanocytes demonstrates that the fitness advantage of cancer-driver mutations can be leveraged to generate progressive multimutant models from primary, differentiated human cells. Overall, we generated melanocytes with up to five precise mutations in key melanoma pathways, opening the way to investigation of genotype-to-phenotype relationships during cancer development.

Consecutive mutations produce ordered progression through expression space in vitro

We related mutations to their expression consequences by profiling cells of each genotype in the editing tree with single-cell RNA sequencing (scRNA-Seq; Fig. 2A) (*30*). We found that as cells harbored increasing numbers of mutations, they continuously progressed in expression space. We profiled cells in multiplex by labeling cells of each genotype with a distinct, DNA-barcoded, cell-surface–protein antibody (cell hashing) (*35*), followed by pooling of all genotypes to assay all the cells in one batch. We retained 11,042 high-quality cell profiles, with a median of 999 cells per genotype (range: 836 to 2360) (*30*). Genotype-agnostic, unsupervised embedding of the profiles into a two-dimensional (2D) space with uniform manifold approximation and projection (UMAP) followed the topology of the melanocyte editing tree (Fig. 2B). WT cell profiles were embedded next to C cell profiles, which were in turn adjacent to CB cells—all forming one continuum, with partial overlap between genotypes (Fig. 2B). The CBT cells, although still adjacent to the CB cells, were connected through a narrow transition and were primarily embedded in a separate cluster of cells that included only genotypes with replicative immortality. Although the CBT3 and CBTP cells were located on either side of the CBT cells, the CBTA melanocytes largely diverged into their own isolated cluster. Finally, CBTP3 and CBTPA cells mapped near their parental CBTP cells (in an overlapping manner), rather than close to the CBT3 or CBTA cells with which they share the 3 and A mutations. These results suggest that as melanocytes acquire sequential cancer-associated mutations, they follow an ordered progression through expression space.

The progression of mutant melanocytes through expression space coincided with modulation of expression programs associated with distinct biological processes. We decomposed the expression profiles of all single cells jointly into expression programs—learned de novo—through consensus nonnegative matrix factorization (cNMF) (Fig. 2C and figs. S4 and S5) (*30*, *36*). The seven programs were used by cells across multiple genotypes, capturing both the continuity of the transitions and shared features between distant genotypes (Fig. 2C). We annotated each program by its top associated genes through manual review and gene set enrichment as "melanocyte," "interferon/p53," epithelial-mesenchymal transition ("EMT"), "Myc/mTORC1," "Myc/mTORC1/Ox-Phos," and cell-cycle–related "S Phase" and "G2/mitosis" (Fig. 2D and tables S3 and S4) (*30*). The melanocyte program was associated with melanocyte lineage genes such as *DCT*, *RAB32*, *TYRP1*, *TRPM1*, *MITF*, and *MLANA*. The highest usage of the melanocyte program was observed in WT melanocytes, whereas it gradually declined in the early mutant genotypes and was still expressed—albeit at lower levels and in fewer cells—in all of the quadruple- and quintuple-mutant melanocytes except the CBT3 melanocytes. The interferon/p53 program was first activated in C melanocytes, reached its apogee in CB melanocytes, and was turned off in CBT melanocytes, such that activation of telomerase through the T mutation led to a sharp decrease in program usage. Both interferon and p53 have been associated with senescence (*37*–*39*), which we observed in CB cells in vitro. This pattern is consistent with cells undergoing stress and telomeric crisis as they age, with telomerase activation reversing these stressors. The EMT program was associated with genes related to invasive potential (e.g., *SERPINE2*, *TIMP3*, *FN1*, *VIM*, *PMEPA1*, *LGALS1*) and was strongly activated in CBT melanocytes, notable for the reported link between *TERT* and EMT (*40*, *41*). The program was also active in those CB cells that were at the phenotypic transition, as well as some of the CBT3 and CBTP cells, particularly those adjacent to CBT cells. EMT has mostly been studied in epithelial cells, and it is unclear how it relates to cell motility and metastatic capacity in tumors with neural crest origins, such as melanomas (*42*, *43*). The Myc/mTORC1 program was activated in CBTA cells (and in some of the CBT3 cells) whereas the Myc/mTORC1/Ox-Phos program was activated in *PTEN* mutant cells (CBTP, CBTP3, and CBTPA). Both cell cycle programs were used at a higher level in all the genotypes that included the T mutation and thus possessed replicative immortality (fig. S6), whereas among the cells without the T mutation, noncycling "G0" C and CB cells gradually moved away from the more distinct G0 WT cells (fig. S6D). Notably, many of the programs also matched those observed in scRNA-seq of human melanoma cell lines (*44*) (fig. S7). Overall, these results demonstrate that melanoma-associated mutation combinations activate and repress specific expression programs that are shared across

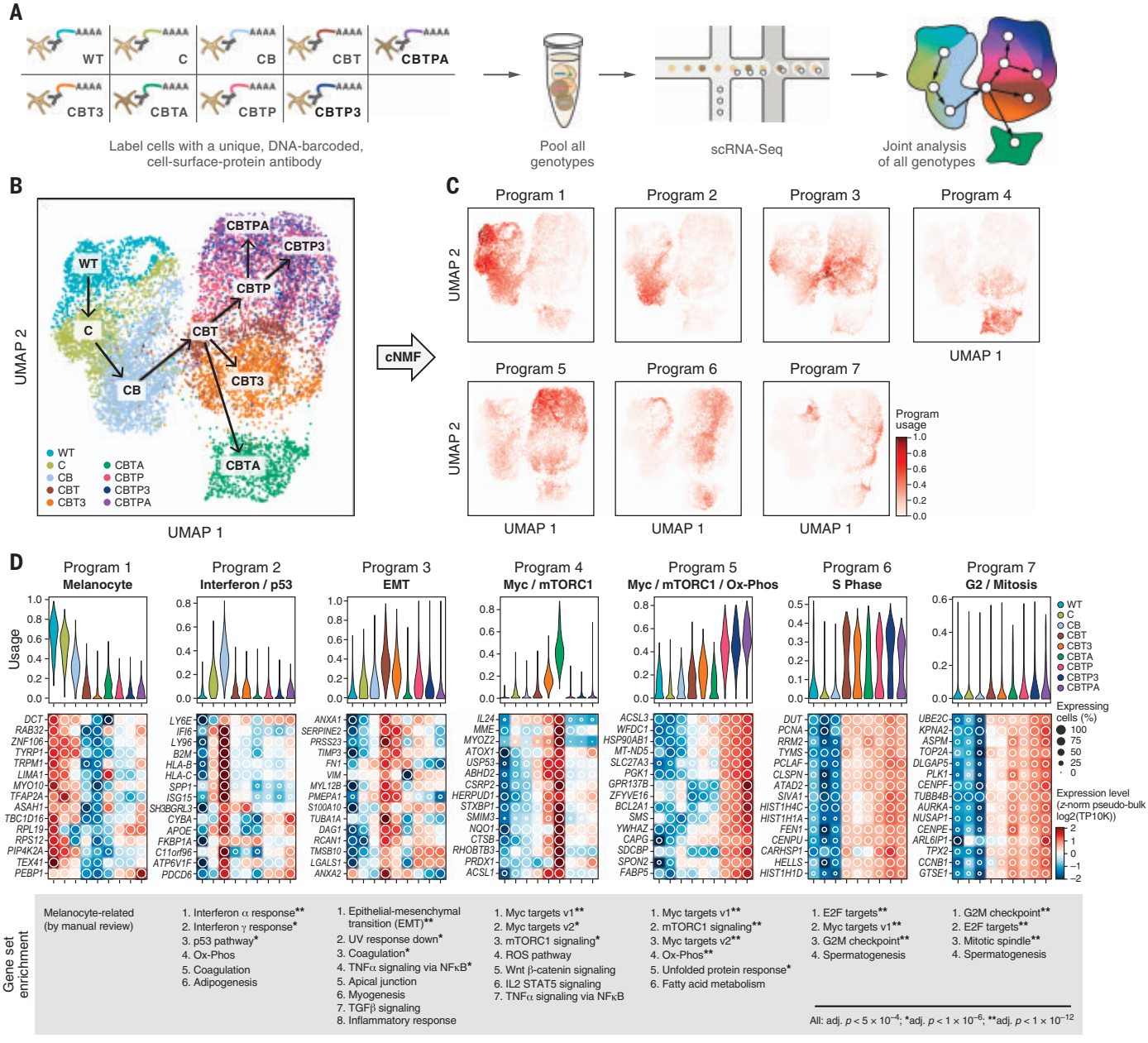


Fig. 2. Consecutive mutations produce ordered progression through expression space and activate shared expression programs. (**A**) Experimental overview to profile gene expression in parallel from cells from nine engineered genotypes with hashed scRNA-seq. (**B**) Gradual progression of cell states with genotype. UMAP embedding of melanocyte scRNA-seq profiles (dots) colored and labeled (boxes) by genotype (legend). Arrows follow the editing tree (as in Fig. 1B). (**C**) Expression programs. UMAP embedding as in (B) colored by per-cell relative usage (color bar) of each of the seven expression programs identified by cNMF. (**D**) Programs reflect key processes and vary across genotypes. Top: Distribution of relative program usages (*y* axis) in single cells of each genotype (*x* axis, color legend). Middle: Aggregate (pseudobulk) expression [Z score of expression (log_2 of transcripts per 10,000 reads, TP10K), color bar] and percent of expressing single cells (white circles) of the 15 top program-associated genes (rows) per genotype (columns). Bottom: Ranked lists of gene sets (MSigDB hallmark) (*83*) enriched in each program (Mann-Whitney *U* test, false discovery rate (FDR) < 5×10^{-4}; *FDR < 10^{-6}; **FDR < 10^{-12}).

genotypes and, in some cases, help explain the overall cellular phenotypes.

Mutation combinations confer distinct, disease-relevant tumor phenotypes in vivo

We next related multimutant genotypes to their tumor phenotypes in vivo by injecting each immortalized melanoma model into the dermis of immunodeficient mice and assessing for disease-relevant features (Fig. 3A). CBT melanocytes were malignant in vivo and formed slow-growing tumors in xenograft models. Between 67 and 111 days, no primary tumor growth was detected [Fig. 3, B and C, ($n = 4$) and ($n = 8$), respectively, black curves; and fig. S8 ($n = 8$)]. However, upon tissue harvest, small nodules were visible at the injection sites. Histologic and immunophenotypic evaluation confirmed the presence of melanoma cells in these nodules (six of six tumors examined; figs. S9 to S12 and tables S5 to S7), with half of the nodules also displaying adjacent features of a congenital nevus (three of six tumors). Over a longer time course (≥150 days), a small tumor (≤14 mm^3)

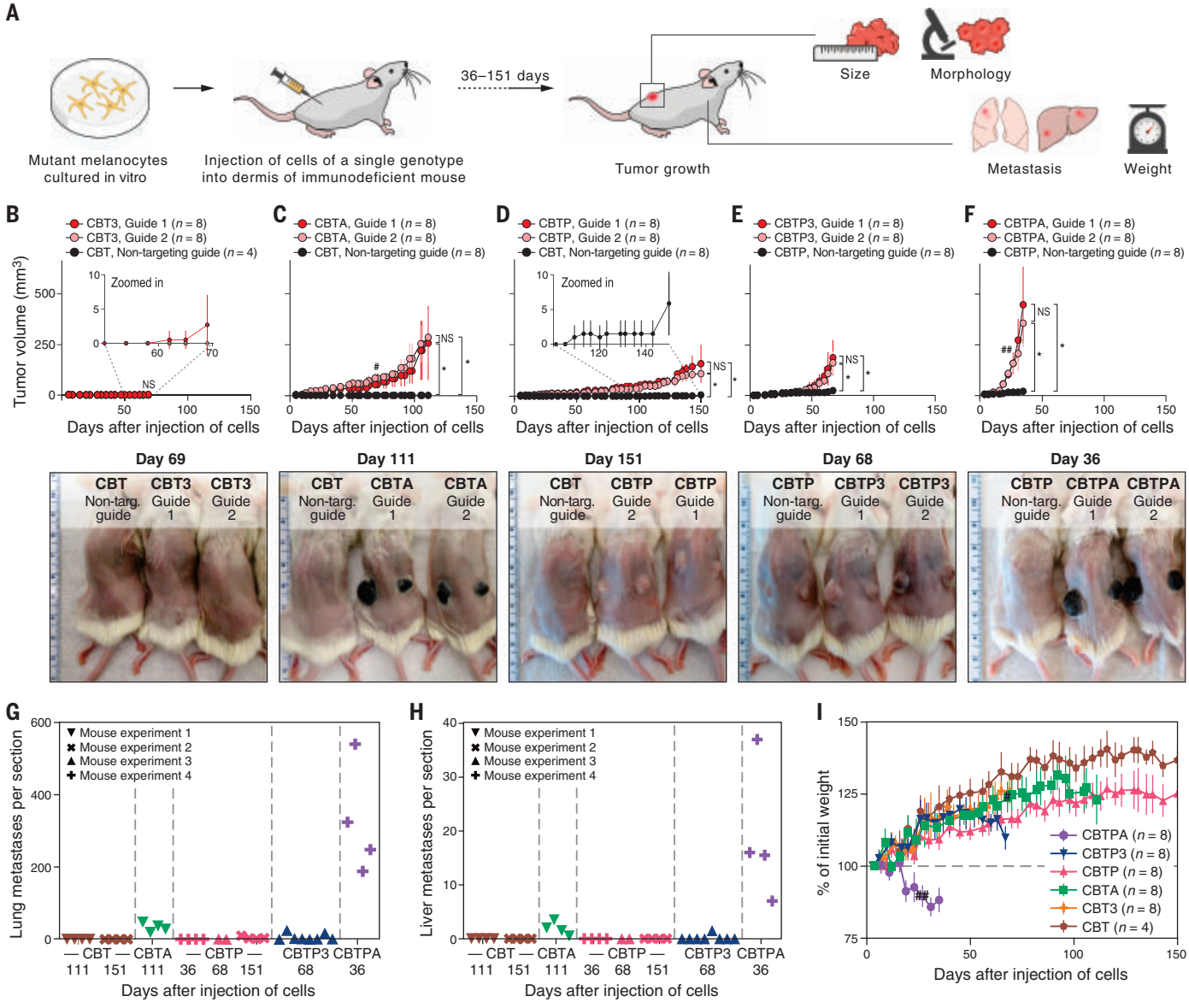


Fig. 3. Mutation combinations confer diverse, disease-relevant phenotypes in vivo. (**A**) Experimental approach to identify disease-relevant phenotypes caused by engineered mutations in vivo. (**B** to **F**) Primary tumor growth of xenografted mutant melanocytes in NSG mice compared with CBT or CBTP control parental cells (as shown) that received nontargeting Cas9 RNP: (B) CBT3, (C) CBTA, (D) CBTP, (E) CBTP3, and (F) CBTPA cells. Top panels: tumor size (cubic millimeters, *y* axis) over time (days, *x* axis) after two intradermal injections, one in each flank. n, number of tumors. Bottom panels: representative images of (shaved) mice harboring mutant cells as marked. A ruler with large, numbered marks in centimeters is shown for scale. (**G** and **H**) Loss of *APC* promotes frequent distant metastases. Average number of individual metastatic foci per section (symbols) of lung (G) or liver (H) tissue in a histologic slide (*y* axis, counted manually) obtained from a single mouse injected with a mutant cell line (genotype indicated by color) and collected after the indicated number of days (*x* axis). Each slide had an average of three lung sections and two liver sections, each from a different lobe. (**I**) Injected CBTPA melanocytes cause rapid weight loss in mice. Percent of initial mouse weight [*y* axis, determined after subtracting primary tumor weights (estimated at 1 g/cm^3) from measured mouse weight]) over time (*x* axis, days). n, number of mice. Data in (G and H) are from the four independent experiments in (C to F). #, two CBTA mice, one from each guide group, were sacrificed for histological inspection (C and I). ##, one CBTPA mouse was euthanized as a result of primary tumor ulceration (F and I). *P < 0.01; ns, not significant; two-tailed, two-sample Student's t test.

occasionally became apparent at the injection site before tissue harvest [7 of 12 injections; Fig. 3D, (n = 8), inset: slight uptick of black curve by day 151; and fig. S13, no growth by day 189 (n = 4)]. These small melanocytic tumors were malignant by dermatopathologic evaluation (four of four tumors examined; fig. S14 and table S8). Thus, in line with observations in patient melanomas (*18*), melanocytes with common melanoma mutations in the endogenous loci of *CDKN2A*, *BRAF*, and *TERT* displayed phenotypic characteristics of early melanomas.

Three alternative quadruple-mutation branches of the editing tree—CBT3, CBTA, and CBTP—each had distinct effects on melanoma phenotypes. CBT3 cells did not produce visible tumors over a period of ~60 days, although by day 69 a few injection sites (3 of 16) began to show small tumors (≤14 mm^3; Fig. 3B: slight uptick of red curve by day 69). By contrast, CBTA cells formed darkly pigmented,

macular (flat) growths by day 10 that advanced to slow-growing, darkly pigmented tumors by day 29 (Fig. 3C). Finally, CBTP melanocytes formed amelanotic tumors in mice that grew faster than CBT3 tumors but slower than CBTA tumors (compare Fig. 3D with Fig. 3B). In all examined quadruple-mutant tumors, histologic and immunophenotypic features of the xenografted melanoma models resembled those of patient melanomas (three of three, CBT3; three of three, CBTA; four of four, CBTP; figs. S15 to S17 and tables S6 to S8). Our findings suggest that in the setting of mutant *CDKN2A*, *BRAF*, and *TERT*, loss of *APC* causes more potent progression of human melanomas than does loss of either of the more commonly mutated genes, *PTEN* or *TP53*.

A fifth engineered mutation led to features of aggressive melanocytic disease. Tumors formed by CBTP3 melanocytes showed a beyond-additive, increased growth rate compared with both CBTP and CBT3 melanocytes, highlighting synergy between the *PTEN* and *TP53* mutations (Fig. 3E, compare with Fig. 3, B and D, at days 67 to 70, $P = 1 \times 10^{-6}$ by two-sided, modified Welch's *t* test) (*30*). CBTP3 tumors also showed evidence of emerging tumor heterogeneity with tumors that were mostly amelanotic (similar to CBTP tumors). However, these frequently had contiguous sectors of dark pigmentation of varying size (fig. S18, see below for analysis of expression and genetic heterogeneity of these tumors). Tumors formed by CBTPA melanocytes had the fastest growth rate of all the engineered melanoma models (Fig. 3F), with mice that had received the CBTPA melanocytes requiring euthanasia by day 36 as a result of primary tumor burden. Similar to CBTA tumors but in contrast to CBTP tumors, CBTPA tumors were uniformly darkly pigmented (Fig. 3F and fig. S18C). This is in line with phenotypes of deep penetrating nevi (DPN) and DPN-like melanomas, in which Wnt pathway mutations have been associated with increased pigmentation (*34*). Both quintuple-mutant genotypes resembled patient melanomas by histologic and immunophenotypic features (four of four, CBTP3; four of four, CBTPA; figs. S19 and S20 and tables S9 and S10).

Metastatic propensity was also associated with tumor genetics. CBTP tumors yielded a small number of lung metastases by day 151 (Fig. 3G), whereas CBTA cells metastasized to both the lung and liver (two common sites of melanoma metastasis) by day 111 (Fig. 3, G and H), as well as to other organs (fig. S21). Tumors formed by CBTPA melanocytes readily metastasized to visceral organs, with numerous metastases visible in the lungs and liver by day 36 (Fig. 3, G and H, and fig. S22). Rapid-onset weight loss, another characteristic of aggressive disease, was also apparent by day 19 after xenograft injection (Fig. 3I). Together with our observations of metastasis in the CBTA model (Fig. 3, G and H, and fig. S21), our findings point to loss of *APC* as a major cause of metastatic disease in this genetic context. This is likely attributable to Wnt pathway activation, although its role in melanoma metastasis has been an open question (*26*, *45*–*47*). The CBTPA melanocyte model also shows that as few as five mutant genes are sufficient to produce aggressive, metastatic human melanoma, at least in an immunodeficient host.

To test for the possibility that additional mutations had accrued during the process of engineering and cellular proliferation, we sequenced the genome of a CBTPA tumor and compared it with the parental WT melanocyte genome to identify somatic events. We did not find mutations of apparent in vivo phenotypic consequence beyond those we had introduced. Notably, we did identify a clonal, two-fold tandem duplication of the melanocyte lineage transcription factor *MITF* (table S11 and fig. S23) (*30*)—a gene amplified in 5 to 10% of melanomas (*8*, *24*, *48*)—but it had no major observed phenotypic consequences (*30*). The spontaneous duplication of a gene frequently amplified in melanomas underscores the similarity of our cell models to melanomas arising in humans. We identified no further somatic alterations of known cancer association, with no additional chromosomal segment amplifications or deletions (fig. S23), only 12 clonal, nonsilent somatic point mutations (not including engineered mutations; table S12 and figs. S24 to S28), and only one structural variant (deletion of *RIC8B*; table S11). These findings reduce the possibility that spontaneous, unplanned mutations are responsible for the phenotypic differences observed between our engineered model genotypes; however, in the absence of deep sequencing of all model genotypes at the time of their phenotypic characterization, contributions from such mutations cannot fully be excluded.

Taken together, our results establish causal relationships between disease characteristics and six different combinations of melanoma mutations in human melanocytes. These results also demonstrate that genome-edited melanocytes recapitulate important aspects of tumor development in vivo.

Genotype-driven intrinsic tumor cell expression programs in vivo

To assess the in vivo cellular phenotypes caused by melanoma mutation combinations, we next performed scRNA-Seq on tumors from our xenografts (Fig. 4A) (*30*). Because each tumor consists of an intricate ecosystem of melanocytic tumor cells (of human origin) along with stromal and immune cells within the tumor microenvironment (of mouse origin), we investigated the effects of mutations on different cell types separately. We first computationally distinguished tumor cells as those cells whose sequencing reads predominately mapped to the human genome (fig. S29) (*30*). For tumor cell analysis, we retained 26,964 high-quality tumor cells, with a median of 1609 (range: 31 to 4999) cells per sample across three replicate tumor samples of each of the CBTP, CBTA, CBTP3, and CBTPA tumors grown in mice for ~1 to 2 months and two replicate tumors for CBTP tumors grown for ~6 months (fig. S30 and table S13) (*30*).

The expression profiles of mutant melanocytes isolated from in vivo tumors were grouped predominantly by genotype. In a genotype-agnostic, unsupervised 2D UMAP embedding of the profiles, melanocytes from CBTA, CBTPA, and 2-month-old CBTP tumors formed one cluster per tumor (Fig. 4B). Melanocytes from CBTP3 tumors formed two clusters per tumor, and melanocytes from the 6-month-old CBTP tumors formed either two or three clusters per tumor (Fig. 4B). In addition, hierarchical clustering of the UMAP cell clusters by the Pearson correlations of their mean expression profiles showed that cell clusters from CBTA and CBTPA tumors were mostly grouped by genotype, whereas 6-month-old CBTP and CBTP3 tumors were partitioned into two clades, one of which was particularly distinct (fig. S31). To assess whether spontaneous genetic changes could be driving the emerging within-sample clusters, we inferred chromosomal copy number alterations (CNAs) from single-cell expression profiles, as previously demonstrated for human tumors (*49*) (fig. S32). If CNA acquisition was the initiating event for formation of supernumerary within-sample clusters, then each supernumerary cluster should demonstrate distinct and clonal CNAs. This was the case for CBTP3 tumor clusters (fig. S32), but not for the within-sample clusters of 6-month-old CBTP tumors, which did not show CNAs (fig. S32). [Note that most of the CBTP3-specific CNAs were not common patient-derived melanoma CNAs, though they sometimes included genes often mutated in melanoma patients) (*16*, *24*, *27*).] Thus, although engineered genotype was the main driver of expression differences between models in vivo, subpopulations of expression states emerged either with time (6-month-old CBTP tumors) or more rapidly as a consequence of additional genotype changes (CBTP3 tumors).

We identified eight expression programs active in mutant melanocytes in vivo and used by cells across multiple genotypes (Fig. 4C and fig. S33) (*30*, *36*), which we annotated as "ribosomal," "Ox-Phos," "interferon/TGFβ," "EMT," "β-catenin/*MITF*," "interferon/TNFα/hypoxia," "protein secretion," and "cell cycle" on the basis of the top genes and gene sets associated with each (Fig. 4D and tables S14 and S15) (*30*). The programs Ox-Phos and β-catenin/

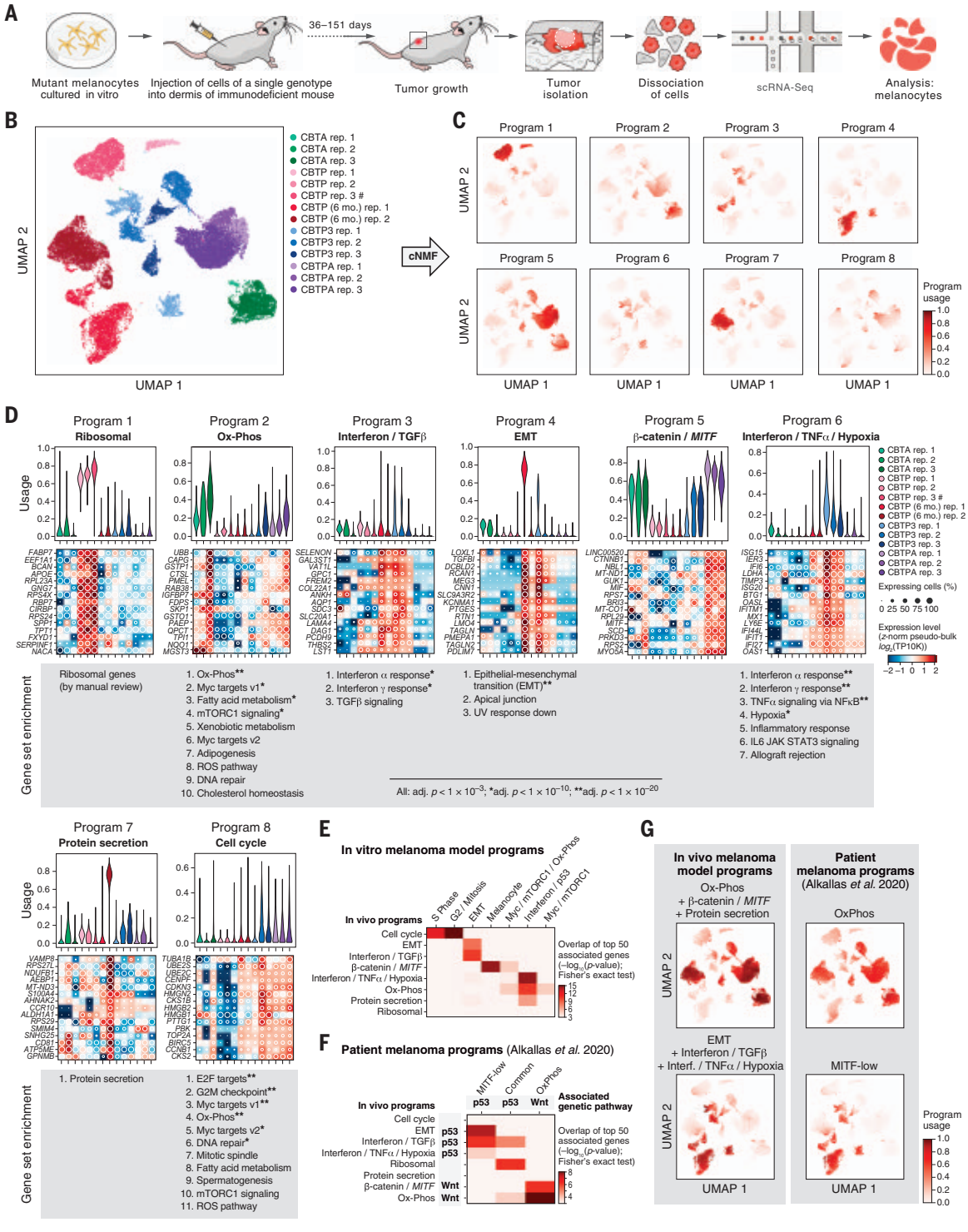


Fig. 4. Genome-edited melanocytic tumors share expression programs with patient melanomas, with matching genetic associations. (**A**) Experimental approach to profile gene expression from tumor cells from xenografts with scRNA-seq. (**B**) Intra- and intergenotype variation of cancer cell states in vivo. UMAP embedding of engineered melanocyte scRNA-seq profiles (dots) colored and labeled (boxes) by genotype and replicate (legend). #, CBTP rep. 3 is a mixture of four tumors from two mice; all other replicates are from a single tumor. (**C**) Expression programs. UMAP embedding as in (B), colored by per-cell relative usages (color bar) of each of the seven expression programs identified by cNMF. (**D**) Programs reflect key cellular processes that vary in usage across genotypes. Top: Distribution of relative program usage (*y* axis) in single cells of each genotype (*x* axis, color legend). Middle: Aggregate (pseudobulk) expression level [Z score of expression level ($\log_2$(TP10K)), color bar] and percent of expressing single cells (white circles) of the 15 top program-associated genes (rows) per genotype (columns). Bottom: Ranked lists of gene sets (MSigDB hallmark) (*83*) enriched in each program (Mann-Whitney *U* test, FDR< 10^{-3}; *FDR < 10^{-10}; **FDR < 10^{-20}). (**E**) Correspondence of in vivo and in vitro programs. Significance of overlap [-$\log_{10}$(*P* value), Fisher's exact test, colorbar] of top 50 associated genes between in vivo (rows) and in vitro (columns, as in Fig. 2) programs. Only overlaps with *P* value < 10^{-3} are shown to account for multiple hypothesis testing. (**F**) Correspondence of in vivo programs and programs in patient melanomas (*16*). Significance of overlap [-$\log_{10}$(*P* value), Fisher's exact test, colorbar] of top 50 associated genes between in vivo (rows) and patient (column) programs. Only overlaps with *P* value < 10^{-3} are shown to account for multiple hypothesis testing. Associations of expression programs with either p53 or Wnt pathway gene mutations are noted. (**G**) Similar usage of melanoma (*16*) and in vivo model programs across in vivo melanoma model single-cell profiles. UMAP embedding [as in (B)], colored by per-cell relative usage of patient melanoma expression programs (right) or sums of relative usages of in vivo melanocyte expression programs (left).

MITF [also associated with *TRPM1*, *CCND1*, *TYRP1*, *MLANA*, and *CDH1* (E-cadherin)] were mostly used by tumors with Wnt pathway activation (*APC* mutation). The interferon/TGFβ, the interferon/TNFα/hypoxia, and the EMT programs were associated with tumors with p53 pathway inactivation (*TP53* mutation). Furthermore, the interferon/TGFβ program predominantly characterized cells from a single within-sample cluster in each of the CBTP3 and 6-month-old CBTP tumors (described earlier; with only one CBTP3 cluster having the aforementioned clonal CNAs; Fig. 4, B and C; program 3), suggesting that these clusters reflect a shared inflammatory expression state, despite their tumor and genotype differences. Most in vivo programs were similar to one or a combination of programs observed during in vitro culture, as judged by the overlap of top associated genes (Figs. 2, C and D, and 4E), consistent with a largely cell-intrinsic origin. However, incomplete overlaps and differences between program usages may also reflect interactions with the microenvironment.

Genetically linked tumor expression programs are shared with patient melanoma tumors with matching genetic associations

Expression programs identified in mutant melanocytes in vivo matched known programs seen in patient melanomas with corresponding genetic associations. Previous studies identified three cancer cell gene expression programs in bulk mRNA expression profiles of patient melanomas (*16*): (i) an MITF-low program (*24*, *50*) associated with low expression of *MITF*; (ii) an OxPhos program associated with genes involved in oxidative phosphorylation and differentiated melanocytes and with low levels of hypoxia-associated genes, and (iii) a less well-defined Common program shared by MITF-low and OxPhos tumors and associated with increased expression of *MITF* and interferon signaling genes. The Ox-Phos and β-catenin/*MITF* in vivo programs (highly used by the darkly pigmented CBTA and CBTPA tumors) showed substantial overlap of top genes with the OxPhos program from patient melanomas, which was itself enriched in patient tumors with mutations in Wnt pathway genes and high levels of pigmentation (Fig. 4F, Fisher's exact test, top 50 associated genes, $P = 1 \times 10^{-15}$ and 7×10^{-7}, respectively) (*16*), thus matching the molecular program, its genetic association, and tumor phenotype. Similarly, the EMT, interferon/TGFβ, and interferon/TNFα/hypoxia in vivo programs (highly used by CBTP3 and 6-month-old CBTP tumors) shared top genes with the patient melanoma MITF-low program ($P = 4 \times 10^{-8}$, 7×10^{-7}, and 2×10^{-4}), which was itself enriched in patient tumors with *TP53* mutations (Fig. 4F) (*16*). Finally, the ribosomal in vivo program—mostly used by 2-month-old CBTP tumors and not observed in vitro—overlapped with the Common program from patient melanomas ($P = 7 \times 10^{-7}$), suggesting that it captures expression features present in patient samples but also shares top genes with interferon/TGFβ and Ox-Phos ($P = 1 \times 10^{-5}$ and 2×10^{-4}), two in vivo programs that overlapped both MITF-low and OxPhos patient melanoma programs, respectively (Fig. 4F).

Moreover, in engineered melanocyte cells from in vivo tumors, the individual activities of the patient melanoma programs OxPhos and MITF-low (derived from bulk tumors) closely matched those of the combined in vivo single-cell programs: (i) Ox-Phos, β-catenin/*MITF* and protein secretion, and (ii) EMT, interferon/TGFβ, and interferon/TNFα/hypoxia, respectively (Fig. 4G). Furthermore, usage of the OxPhos and MITF-low patient melanoma programs in our models showed intratumoral heterogeneity (Fig. 4G, right column), as has been observed in scRNA-seq of patient melanomas (with programs comparable to those identified here, fig. S34) and in cell lines (*49*). Overall, our engineered melanocyte models recapitulate the expression states and genetic associations in patient melanomas and suggest that the expression programs described to date in patient melanomas are a composite of coincident, biologically distinct programs that are not merely associated with specific gene mutations but rather are caused by them, albeit with intratumoral variation.

Tumor genotypes shape the composition and expression state of infiltrating stromal and immune cells

Next, we estimated the effect of malignant cell genotype on the tumor microenvironment by analyzing the scRNA-seq profiles of 13,332 mouse cells (median cells per sample: 576, range: 88 to 4043) across all tumors (Fig. 5A) (*30*). The cells spanned immune cell types (neutrophils, dendritic cells, plasmacytoid dendritic cells, and M1 and M2 macrophages), as well as endothelial and epithelial cells, pericytes, and cancer-associated fibroblasts (Fig. 5B and fig. S35), which we annotated by marker gene expression (figs. S36 to S39) (*30*). [T, B, and natural killer (NK) cells are absent, as expected in these immunodeficient NOD. Cg-*Prkdc*^{scid} *Il2rg*^{tm1Wjl}/SzJ (NSG) mice.] As observed in patient tumors (*49*, *51*), profiles from cells of the microenvironment are primarily grouped by cell type (Fig. 5B). We observed tumor age-related changes in cell composition among cancer-associated fibroblasts, which shift from more balanced ratios of contractile and immunomodulatory expression phenotype populations in 2-month-old tumors to 97 to 98% of fibroblasts having the contractile phenotype in 6-month-old CBTP tumors (Fig. 5C), in line with previous observations (*52*).

Melanocyte genotype altered the cellular composition of the tumor microenvironment (Fig. 5C)—most notably neutrophils, which are known to be associated with poor early stage melanoma prognosis (*53*). Neutrophils constituted ~40% of the tumor microenvironment cells in CBTA and CBTP3 tumors on average but were nearly absent from CBTP tumors of any age and constituted only ~2% of the microenvironment cells in CBTPA tumors (which are also the tumors grown in mice for the shortest time because of their fast growth rate) [Fig. 5, B and C, and figs. S30, S35B, and S40; family-wise sign error rate (Bayesian proxy for family-wise error rate controlled P value) < 0.01 for comparisons of CBTA or CBTP3 versus CBTP (2 or 6 months), hierarchical Bayesian multinomial logistic mixed effects model (*30*)]. The differences in neutrophil infiltration between tumors sharing many of the same mutant genes underscore the importance of mutation combinations in shaping the tumor microenvironment. We hypothesize that neutrophil infiltration in individual tumors may be due to differences in tumor-immune cell communication. Supporting this hypothesis, the melanocytes from the two CBTA tumors and one CBTP3 tumor with highest neutrophil infiltration [CBTA replicates 1 and 2 (rep. 1 and 2); CBTP3 rep. 1; fig. S35B] expressed the chemoattractant *CCL2*, known to attract and activate neutrophils (fig. S41A) (*54*). Notably, neutrophils in CBTA and CBTP3 tumors also displayed shifts in distribution across different cell states, partly tracking with genotype (Fig. 5, D to F, and fig. S35). Neutrophils from the two CBTA tumors with the highest neutrophilic infiltrate were associated with an expression program previously observed in tumor-infiltrating and tumor-promoting neutrophils (N5), whereas CBTP3-infiltrating neutrophils in the most enriched replicate (rep. 1), expressed programs that resembled the expression state of healthy-tissue neutrophils (N1 and N3) (*55*) (Fig. 5, D to F, and figs. S35 and S41B).

Melanocyte genotype also influenced the cellular state of other immune cells in the tumor microenvironment. Examining each cell type for genotype-associated expression differences, we observed that macrophages from different tumor genotypes grouped in genotype-related patterns (Fig. 5G). Although M2 macrophages—most prominent in CBTA tumors—did not show clear tumor genotype-specific expression changes, M1 macrophages preferentially activated three different expression programs depending on tumor genotype (Fig. 5, H and I, fig. S41C, and tables S16 and S17): (i) a "complement/ribosomal" program, enriched in ribosomal protein genes and a subset of complement component genes (*C1qa*, *C1qb*, *C1qc*) that were reported to increase in response to apoptotic cells and

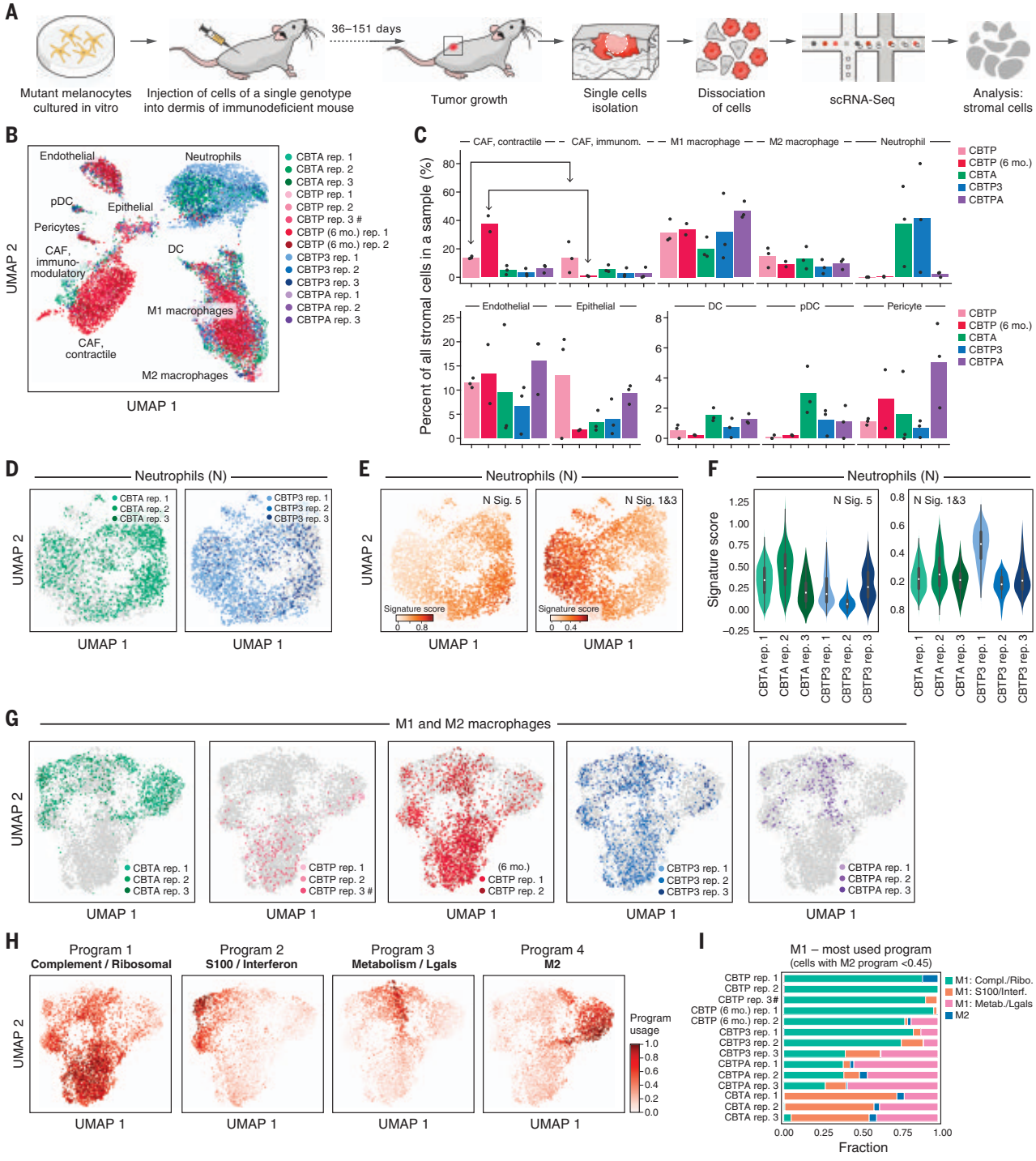


Fig. 5. Tumor genotype shapes tumor microenvironment composition. (**A**) Experimental approach to profile gene expression from mouse cells in tumor xenografts with scRNA-seq. (**B** and **C**) Remodeling of the tumor microenvironment cellular composition by cancer cell genotype and duration in mice. (B) UMAP embedding of tumor microenvironment scRNA-seq profiles (dots) colored by genotype and replicate (legend) and labeled by cell type. #, CBTP rep. 3 is a mixture of four tumors from two mice, whereas all other replicates are from a single tumor. (C) Mean (bar) and individual (dots) percent (*y* axis) of tumor microenvironment cells of each type (*x* axis) in each genotype (color). (**D** to **F**) Diversity of neutrophil expression programs in tumors of different genotypes. (D) UMAP embedding of neutrophil single-cell profiles (dots) from specific tumor genotypes (colored) and all other genotypes (gray). (E) UMAP embedding of neutrophil profiles (as in D) highlighting only the neutrophils in CBTA and CBPT3 tumors colored by per-cell score of neutrophil expression signatures, N5 (associated with tumor growth), and N1N3 (more similar to circulating and healthy-tissue neutrophils); signatures previously described (*55*). (F) Distribution of per-cell neutrophil expression program score (*y* axis) in neutrophils from CBTA (green) and CBTP3 (blue) tumors. *P* value < 0.001 (N5), *P* value < 0.001 (N1N3), Kruskal-Wallis rank-sum test, degrees of freedom = 5. (**G** to **I**) Effect of tumor genotype on macrophage expression programs. (G and H) UMAP embedding of M1 and M2 macrophage single-cell profiles (dots), colored by specific tumor genotypes (G) or by per-cell relative usage of macrophage gene expression cNMF programs (H). (I) Fraction of cells (*x* axis) with the highest score in each of four cNMF programs (colors) among M1 cells (defined as cells with an M2-related program score <0.45) in each replicate (*y* axis).

during macrophage polarization toward the M2 phenotype (*56*, *57*); (ii) an "S100/Interferon" program, enriched in genes related to TNFα response through NF-κB (e.g., *Cebpb* and *Atf3*), interferon response (e.g., *Ifitm1* and *Ifitm3*), and S100 calcium-binding cytosolic proteins (e.g., *S100a4/6/10/11*); and (iii) a "metabolism/Lgals" program enriched in genes involved in cellular metabolism (e.g., *Aldoa* and *Ldha)* and cell-cell interactions such as *Lgals1* or *Lgals3*, also involved in immunoregulatory functions (*58*). CBTP tumor M1 macrophages primarily expressed the complement/ribosomal program (Fig. 5I), whereas M1 macrophages from the CBTPA tumors mostly expressed the metabolism/Lgals program. Lastly, the CBTP3 tumors had macrophage distributions matching either the CBTP (CBTP3 rep. 1 and 2) or CBTPA (CBTP3 rep. 3) patterns (Fig. 5I), perhaps because of genetic heterogeneity (CNAs; fig. S32) (*30*) seen between CBTP3 tumors. Finally, M1 macrophages from the CBTA tumors predominantly expressed either the S100/interferon program or the metabolism/Lgals program. M2-scoring cells lie in proximity to M1-like cells that score highly for either the complement/ribosomal or metabolism/Lgals programs, suggesting possible alternative paths from M1 to M2 in tumors from different genotypes, depending on which of these two programs are used. Overall, these results show that tumor mutation combinations shape not only the cellular composition of the tumor microenvironment but also the cellular state of the individual cell types within the microenvironment.

Tumor histological features are genotype specific and coincide with genotype-associated expression programs in patient melanomas

We examined the association between the genotypes of mutant melanocytes and microscopic tumor appearance by assessing whether genotype can be predicted from histopathological images alone, indicating a relationship between the two. To this end, we trained a convolutional neural network model (*59*) on hematoxylin and eosin (H&E)-stained tumor sections of our mutant melanocytes grown as in vivo mouse xenografts, supervised by genotype (Fig. 6A). We then applied the model to predict the probability of each genotype on a per-tile (2048 pixels × 2048 pixels) basis, and combined the predictions (summing per-tile probability vectors in a tumor section and selecting the genotype with maximum probability) to call an overall genotype (CBT, CBT3, CBTA, CBTP, CBTP3, or CBTPA), if there was sufficient prediction certainty (Fig. 6B; entropy of probability vector <0.2) (*30*). We trained on 56 whole–microscope slide images (37%, corresponding to 5533 tiles) and tested on 94 (63%, 16,118 tiles), ensuring that no mouse contributed images to both the training and test sets.

The model classified 76% of sections and had high accuracy [Fig. 6C, area under the curve (AUC) range 0.89 to 1.00, compared with an AUC of 0.50 resulting from the null model, the predictions of which are random] (*30*), with perfect assignment for CBT, CBT3, CBTA, and CBTPA tumor sections (Fig. 6D), possibly reflecting the within-genotype histological homogeneity of these models. The most common misclassification was between CBTP and CBTP3 tumor sections (28% of CBTP sections were classified as CBTP3, and 3% of CBTP3 sections were classified as CBTP), suggesting overlap in histopathological features between these tumor genotypes (Fig. 6D). This may be consistent with the observed expression similarity between these malignant cells in vitro (Fig. 2B) and in vivo (fig. S31). Together, these findings show that mutant melanocyte genotypes give rise to distinguishable tumor histologies.

We then asked whether the distinguishable tumor histologies of our melanoma models were reflected, to any extent, in patient melanomas. To this end, we tested whether the neural network model trained on our genetically distinct xenograft tumors (Fig. 6A) showed any predictive signal on patient melanoma histology. We first used the model to classify histological slides of patient melanomas by our six genome-engineered genotype labels (CBT, CBT3, CBTA, CBTP, CBTP3, or CBTPA) and then grouped all genotype labels containing a given mutation to produce a proxy score (sum of probabilities of genotypes containing the mutation) for loss-of-function mutation status of *APC*, *TP53*, or *PTEN* in patient melanomas (*24*, *30*). Inference of *APC* loss-of-function status resulted in an AUC of 0.58 (*APC* or *CTNNB1* mutations, grouped together to increase the number of tumors). However, the 95% confidence interval (CI) (*30*) included random predictions (0.49 to 0.66; Fig. 6E), and consequently the result was not statistically significant. Inference of *TP53* or *PTEN* mutations approached random prediction [AUC: 0.52, 0.53 respectively, and 95% CI 0.44 to 0.59 and 0.44 to 0.61, respectively; Fig. 6E, mutation annotations as in (*16*)]. Despite lacking statistical significance, the AUC for predicting mutations in the Wnt pathway was within the range of reported results from models trained and tested on patient melanoma histopathology (*APC* AUC: 0.44 to 0.66, *CTNNB1* AUC: 0.52 to 0.64, *TP53* AUC: 0.59 to 0.62, *PTEN* AUC: 0.44 to 0.66) (*60*, *61*).

Because not all patient melanomas with Wnt pathway or *TP53* mutations are readily identifiable as such, we next used expression programs previously associated with Wnt pathway or *TP53* mutations as biomarkers for either the mutations themselves or their functional effects (*16*). The model showed a statistically significant ability to predict expression programs associated with either the Wnt pathway (OxPhos program, AUC: 0.74, 95% CI 0.67 to 0.81) or *TP53* mutations (MITF-low and Common programs, AUC: 0.63, 95% CI 0.56 to 0.70) (Fig. 6E). We corroborated these findings by verifying that our model did not exhibit predictive power on "wrong" or genotype-mismatched labels (e.g., by using *APC* loss-of-function predictions to predict which tumors express expression programs associated with *TP53* mutations) (fig. S42A).

The fact that our model—trained on histopathology of genome-engineered melanocytes grown in mice—can predict genotype-associated expression programs of patient melanomas from histopathology alone, even to a limited extent, is notable given the genetic complexity and heterogeneity of human melanomas. Our results suggest that genotype-associated expression states are reflected in the histopathological features of human melanomas, and that elements of these features are shared between our engineered melanocyte models and melanomas arising in patients.

Discussion

We have shown that the same fitness advantage by which cancer mutations drive clonal expansions in human tumorigenesis can be harnessed to generate multimutant models of cancer from primary human cells in a stepwise manner. By avoiding single-cell cloning, this model building strategy is applicable to cancers arising from differentiated cells, which may not have sufficient replication potential to grow to large populations from a single cell. Because no selection markers are introduced, there is no exogenous DNA that could alter gene regulation or function. As a result, the genome-edited cell models are amenable to selection-marker–based experiments, including comparative molecular studies, genome-wide genetic screens (*62*–*65*), and pooled genetic perturbations, including those with high content readouts (*66*, *67*). Applying such approaches in a human context with matched genetic controls may be of particular use for discovery of therapeutic targets that could escape identification in nonhuman models as a result of interspecies differences, or in human cell lines or patient-derived xenografts because of extensive intersample genetic differences. The phenotypes we report are a product of one particular order of mutations; however, the editing strategy lends itself to modeling alternative sequential ordering of mutation combinations in future studies. Overall, the stepwise, multimutant, genetically precise nature of these human cell models enables study of genotype to phenotype relationships in a context that approximates several defining features of human cancer pathogenesis.

The importance of studying combinations of mutations when linking cancer genotypes

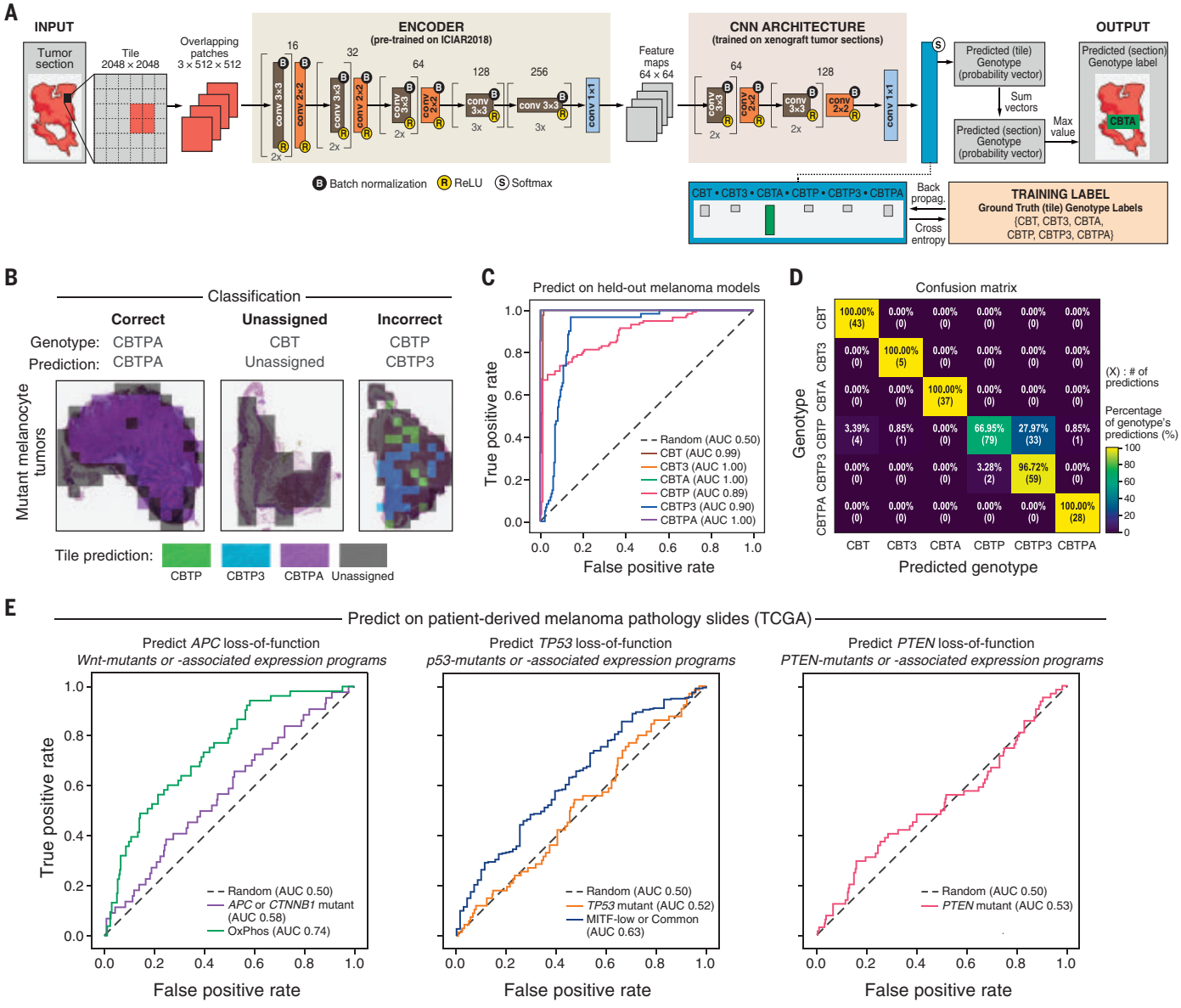


Fig. 6. Tumor genotype leads to distinct histological features that also associate with genotype-linked expression states in patient melanomas. (**A**) Computational approach to classify histological slides into engineered genotypes. (**B**) Test set classification examples. Classification of individual tiles (colored squares overlying tissue images), the aggregated classification for the entire section ("prediction"), and the true genotype ("genotype") for three examples. (**C** and **D**) Successful prediction of genotype from histology in held-out mutant melanocyte in vivo tumor section images. (C) Receiver operating characteristic (ROC) curves of the prediction false positive rate (*x* axis) and true positive rate (*y* axis) at each probability threshold for each genotype (color). The area under the curve (AUC) is indicated for each genotype in the legend. (D) Percentage (color bar) of samples from a given genotype (*y* axis) that received each genotype classification (*x* axis). The percentage and number of such predictions are displayed within each cell. (**E**) Inferring genotype and genotype-associated expression states in patient melanomas (from TCGA) on the basis of images of H&E stained tumor sections. ROC curves were obtained through: (left) predicting *APC* loss-of-function genotype (a Wnt pathway gene) and setting true positive labels to be either Wnt pathway mutants or a Wnt-associated expression program; (middle) predicting *TP53* loss-of-function genotype and setting true positive labels to be *TP53* mutants or *TP53*-associated expression programs; or (right) predicting *PTEN* loss-of-function genotype and setting true positive labels to be *PTEN* mutants (no *PTEN*-associated expression program available).

to phenotypes is highlighted by our observation of how the effect of a given single mutation depends on genetic context. For example, CBTPA and CBTP3 melanocytes in vitro were more similar in expression to CBTP melanocytes than to CBTA or CBT3 melanocytes, suggesting that *PTEN* loss modulates the effect of *APC* and *TP53* loss on gene expression. In another example, a *TP53* mutation produced different effects depending on which other mutations were present: CBT3 melanocytes in vivo did not form sizable primary tumors, but CBTP3 tumors grew faster than CBTP tumors, reflecting an interaction between *PTEN* and *TP53* mutations (*68*) (although no significant association between these two mutations has been noted in patient melanomas) (*16*, *24*, *69*). These nonadditive effects also extend to the tumor microenvironment. For example, the abundance of neutrophils in CBTA and CBTP3 tumors and their absence or near absence in CBTP and CBTPA tumors was not attributable to a difference in a single mutation, as the two tumor groups shared many mutations.

Future studies may explore the effect on therapeutic response of the complex genetic interactions apparent in these melanoma models in cell-intrinsic (tumor cell) and -extrinsic (microenvironment) ways, especially because resistance to targeted therapies often cannot be explained by single-gene mutations (*70*, *71*). Because of genetic epistasis, understanding the mutational landscape of human cancer is a combinatorial problem and its dissection requires modeling strategies that can scale to multiple mutations, such as the one presented in this study, along with appropriate analytics.

Several challenges in melanoma research can be approached by using our genetically precise human models. First, our models can help shed light on the molecular basis of metastasis (*71*). No mutation has been conclusively linked to metastasis in melanomas (*26*, *27*) and the role of the Wnt pathway is debated (*26*, *45*–*47*). Our human models establish a causal link between an activating mutation in the Wnt pathway (through genetic inactivation of *APC*) and metastasis, markedly bolstering complementary lines of evidence for the role of Wnt in melanoma metastasis, most notably two mouse models (*43*, *72*) and genetic evidence from a single patient (*4*). Our observations that *APC* mutations undergo positive selection in culture suggests that mutations driving metastasis may confer a fitness advantage even within the primary tumor, which would explain why metastasis-specific mutations have not been seen in melanomas (*26*, *27*). Moreover, our metastatic melanoma models reproducibly develop spontaneous metastases with short latency, addressing a recognized challenge in animal models in which rapid growth of primary tumors may not allow sufficient time for metastases to develop (*71*). This growth pattern makes CBTA and CBTPA cells tractable systems for studying key intrinsic and extrinsic factors in metastasis, in addition to allowing direct comparison of primary tumor and paired metastases. Although efforts to drug the Wnt pathway in cancer have yet to achieve marked clinical success, our well-characterized and genetically precise human metastatic melanoma models may support efforts to reveal melanoma-specific Wnt vulnerabilities amenable to therapeutic intervention against metastasis, the most lethal aspect of malignancy (*73*).

Another major challenge in melanoma research is understanding response and resistance to targeted therapies (*71*, *74*)—e.g., BRAF and MEK inhibitors (*75*)—in which known genetic mechanisms explain ~60% of resistance (*70*). Although tumor cell gene expression programs are also thought to drive resistance (*76*), models of these programs are needed (*71*, *74*). Our human melanoma models recapitulate the MITF-low expression signature associated with drug resistance (*77*) and demonstrate varying degrees of its presence across different genetic backgrounds (all sharing the *BRAF* V600E activating mutation); our models also link this expression signature to *TP53* inactivation. Our models may therefore be useful in examining the effect of combinations of mutations and their expression states on drug response and resistance. Additionally, precise isogenic models allow for well-controlled chemical and genomic knockout screens that may reveal approaches for disfavoring the resistance-associated gene expression program.

In the context of understanding the immune response and the role of age or environmental exposure (*71*), although our models do not capture the adaptive immune system, we have linked precise melanoma mutations to changes in innate immune cells within a given tumor and provide a genetically controlled system for further study. Future studies may explore the integration of our models into systems that recapitulate human tissue interactions, such as 3D skin (*78*) or skin organoids (*79*), the implantation of our cells into mice of different ages, or the cells' exposure to ultraviolet radiation.

Lastly, the approach we present opens the door for creation of a broader set of human cellular cancer models for melanoma research. For example, the editing tree can be expanded to characterize the molecular and phenotypic consequences of other mutations that are commonly identified in patients but are still poorly understood, such as those in *ARID2* in the SWI/SNF pathway (*8*, *23*). Furthermore, melanocytes from different bodily locations and developmental stages (known to affect melanocyte expression states) (*80*) can be used as alternative starting points, as can melanocytes from donors of both sexes, varying genetic backgrounds, or different skin tones. We opted to engineer neonatal, foreskin-derived melanocytes as they are the most commonly used experimental human melanocytes, but in the future these can be compared with models generated from different types of melanocytes. Finally, the same editing approach can be applied to generate human models of uveal, acral, and mucosal melanomas, diseases for which precise cellular models are still lacking and targeted therapies are still not available (*71*). Though precise engineering of chromosomal amplifications and deletions that typify the latter two subtypes may be more challenging (*81*), an editing approach such as the one described here enables efficient testing of candidate driver genes both individually and in combinations (*71*).

Stepwise genome editing of human primary differentiated cells can convert genetic maps of cancer into genotype-phenotype understanding, as we demonstrated through use of a series of progressive human models of melanoma development. The isogenic cellular models and their associated in vitro and in vivo single-cell expression profiles and histopathology images are a resource spanning multiple melanoma genotypes, including early melanoma precursors that are notoriously difficult to obtain from patient-derived sources (*26*). Genome-edited human models advance knowledge of the genetic basis of human malignancy by ascribing causation of malignant phenotypes to defined sets of genetic alterations, thus allowing for their further study in isogenic human models of disease.

Materials and methods summary

Engineered melanocytes

Genome engineering was performed on primary human epidermal melanocytes derived from the foreskin of a neonatal, lightly pigmented male [ThermoFisher Scientific, Cat. C0025C, donor 1583283; classified as nonhuman subjects research by Broad Institute Office of Research Subject Protection (ORSP-1487)], which were cultured at 37°C, 5% CO_2, and 5% O_2 in M254 medium supplemented with HMGS-2 melanocyte growth supplement, with medium and cells from ThermoFisher Scientific. Cas9 protein, tracrRNA, and guides were purchased from IDT and prepared according to manufacturer instructions. Electroporation was performed with the P3 Primary Cell Nucleofector Kit and Nucleofector 4D System from Lonza. Recombinant AAV was used to deliver a donor DNA template for precision editing.

Mouse xenografts

Female NSG mice (4 to 6 weeks old) received two intradermal injections, one in each flank. Tumor size and body weight were assessed twice per week. All procedures were performed under Massachusetts Institute of Technology Committee for Animal Care protocol 0036-01-15.

Single-cell RNA-Seq

Cells grown in vitro were processed with 10x Genomics Single Cell 3′ v3, with hashing generated as described previously (*82*). For in vivo expression profiles, single cells were dissociated from a cube excised from the center of each xenograft tumor, and cells were processed with 10x Genomics 3′ v2. Expression programs were identified with cNMF as implemented in Kotliar *et al.* (*36*).

Histopathology and machine learning

Paraffin-embedded tumors were sectioned multiple times, stained with H&E, and imaged, yielding 150 whole-slide images (WSIs) from 52 blocks. The WSI were processed to obtain tissue patches, a fraction of which were used as input for a two-stage convolutional neural network based on previous work (*59*). Testing

was performed on nontraining patches and on melanoma patient samples from The Cancer Genome Atlas (TCGA), processed with the same pipeline. Additional and more detailed materials and methods are available in the supplementary materials (*30*).

REFERENCES AND NOTES

1. L. A. Garraway, E. S. Lander, Lessons from the cancer genome. *Cell* **153**, 17–37 (2013). doi: 10.1016/j.cell.2013.03.002; pmid: 23540688
2. B. Vogelstein *et al*., Cancer genome landscapes. *Science* **339**, 1546–1558 (2013). doi: 10.1126/science.1235122; pmid: 23539594
3. S. Turajlic *et al*., Tracking Cancer Evolution Reveals Constrained Routes to Metastases: TRACERx Renal. *Cell* **173**, 581–594.e12 (2018). doi: 10.1016/j.cell.2018.03.057; pmid: 29656895
4. J. Z. Sanborn *et al*., Phylogenetic analyses of melanoma reveal complex patterns of metastatic dissemination. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 10995–11000 (2015). doi: 10.1073/pnas.1508074112; pmid: 26286987
5. J. M. Zaretsky *et al*., Mutations Associated with Acquired Resistance to PD-1 Blockade in Melanoma. *N. Engl. J. Med.* **375**, 819–829 (2016). doi: 10.1056/NEJMoa1604958; pmid: 27433843
6. G. Ciriello *et al*., Emerging landscape of oncogenic signatures across human cancers. *Nat. Genet.* **45**, 1127–1133 (2013). doi: 10.1038/ng.2762; pmid: 24071851
7. M. Imielinski *et al*., Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell* **150**, 1107–1120 (2012). doi: 10.1016/j.cell.2012.08.029; pmid: 22980975
8. E. Hodis *et al*., A landscape of driver mutations in melanoma. *Cell* **150**, 251–263 (2012). doi: 10.1016/j.cell.2012.06.024; pmid: 22817889
9. M. Ghandi *et al*., Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* **569**, 503–508 (2019). doi: 10.1038/s41586-019-1186-3; pmid: 31068700
10. M. J. Garnett *et al*., Systematic identification of genomic markers of drug sensitivity in cancer cells. *Nature* **483**, 570–575 (2012). doi: 10.1038/nature11005; pmid: 22460902
11. J. Barretina *et al*., The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–607 (2012). doi: 10.1038/nature11003; pmid: 22460905
12. L. Cong *et al*., Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013). doi: 10.1126/science.1231143; pmid: 23287718
13. P. Mali *et al*., RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013). doi: 10.1126/science.1232033; pmid: 23287722
14. J. Drost *et al*., Sequential cancer mutations in cultured human intestinal stem cells. *Nature* **521**, 43–47 (2015). doi: 10.1038/nature14415; pmid: 25924068
15. M. Matano *et al*., Modeling colorectal cancer using CRISPR-Cas9-mediated engineering of human intestinal organoids. *Nat. Med.* **21**, 256–262 (2015). doi: 10.1038/nm.3802; pmid: 25706875
16. R. Alkallas *et al*., Multi-omic analysis reveals significantly mutated genes and DDX3X as a sex-specific tumor suppressor in cutaneous melanoma. *Nat. Can.* **1**, 635–652 (2020). doi: 10.1038/s43018-020-0077-8; pmid: 35121978
17. B. C. Bastian, The molecular pathology of melanoma: An integrated taxonomy of melanocytic neoplasia. *Annu. Rev. Pathol.* **9**, 239–271 (2014). doi: 10.1146/annurev-pathol-012513-104658; pmid: 24460190
18. A. H. Shain *et al*., The Genetic Evolution of Melanoma from Precursor Lesions. *N. Engl. J. Med.* **373**, 1926–1936 (2015). doi: 10.1056/NEJMoa1502583; pmid: 26559571
19. D. C. Bennett, Genetics of melanoma progression: The rise and fall of cell senescence. *Pigment Cell Melanoma Res.* **29**, 122–140 (2016). doi: 10.1111/pcmr.12422; pmid: 26386262
20. E. Hodis, L. A. Garraway, in *Melanoma*, D. E. Fisher, B. C. Bastian, Eds. (Springer, 2017), vol. 161, pp. 1–23.
21. F. W. Huang *et al*., Highly recurrent TERT promoter mutations in human melanoma. *Science* **339**, 957–959 (2013). doi: 10.1126/science.1229259; pmid: 23348506
22. S. Horn *et al*., TERT promoter mutations in familial and sporadic melanoma. *Science* **339**, 959–961 (2013). doi: 10.1126/science.1230062; pmid: 23348503
23. M. Krauthammer *et al*., Exome sequencing identifies recurrent somatic RAC1 mutations in melanoma. *Nat. Genet.* **44**, 1006–1014 (2012). doi: 10.1038/ng.2359; pmid: 22842228
24. R. Akbani *et al*., Genomic Classification of Cutaneous Melanoma. *Cell* **161**, 1681–1696 (2015). doi: 10.1016/j.cell.2015.05.044; pmid: 26091043
25. L. B. Alexandrov *et al*., Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013). doi: 10.1038/nature12477; pmid: 23945592
26. A. H. Shain, B. C. Bastian, From melanocytes to melanomas. *Nat. Rev. Cancer* **16**, 345–358 (2016). doi: 10.1038/nrc.2016.37; pmid: 27125352
27. A. H. Shain *et al*., Genomic and Transcriptomic Analysis Reveals Incremental Disruption of Key Signaling Pathways during Melanoma Evolution. *Cancer Cell* **34**, 45–55.e4 (2018). doi: 10.1016/j.ccell.2018.06.005; pmid: 29990500
28. H. Zeng *et al*., Bi-allelic Loss of CDKN2A Initiates Melanoma Invasion via BRN2 Activation. *Cancer Cell* **34**, 56–68.e9 (2018). doi: 10.1016/j.ccell.2018.05.014; pmid: 29990501
29. M. R. Stratton, P. J. Campbell, P. A. Futreal, The cancer genome. *Nature* **458**, 719–724 (2009). doi: 10.1038/nature07943; pmid: 19360079
30. Materials, methods, and supplementary text are available as supplementary materials.
31. D. P. Dever *et al*., CRISPR/Cas9 β-globin gene targeting in human haematopoietic stem cells. *Nature* **539**, 384–389 (2016). doi: 10.1038/nature20134; pmid: 27820943
32. A. Hendel *et al*., Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nat. Biotechnol.* **33**, 985–989 (2015). doi: 10.1038/nbt.3290; pmid: 26121415
33. S. C. Knight *et al*., Dynamics of CRISPR-Cas9 genome interrogation in living cells. *Science* **350**, 823–826 (2015). doi: 10.1126/science.aac6572; pmid: 26564855
34. I. Yeh *et al*., Combined activation of MAP kinase pathway and β-catenin signaling cause deep penetrating nevi. *Nat. Commun.* **8**, 644 (2017). doi: 10.1038/s41467-017-00758-3; pmid: 28935960
35. M. Stoeckius *et al*., Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biol.* **19**, 224 (2018). doi: 10.1186/s13059-018-1603-1; pmid: 30567574
36. D. Kotliar *et al*., Identifying gene expression programs of cell-type identity and cellular activity with single-cell RNA-Seq. *eLife* **8**, e43803 (2019). doi: 10.7554/eLife.43803; pmid: 31282856
37. M. De Cecco *et al*., L1 drives IFN in senescent cells and promotes age-associated inflammation. *Nature* **566**, 73–78 (2019). doi: 10.1038/s41586-018-0784-9; pmid: 30728521
38. S. Wang *et al*., Interferon-γ induces senescence in normal human melanocytes. *PLOS ONE* **9**, e93232 (2014). doi: 10.1371/journal.pone.0093232; pmid: 24681574
39. A. Rufini, P. Tucci, I. Celardo, G. Melino, Senescence and aging: The critical roles of p53. *Oncogene* **32**, 5129–5143 (2013). doi: 10.1038/onc.2012.640; pmid: 23416979
40. Z. Liu *et al*., Telomerase reverse transcriptase promotes epithelial-mesenchymal transition and stem cell-like traits in cancer cells. *Oncogene* **32**, 4203–4213 (2013). doi: 10.1038/onc.2012.441; pmid: 23045275
41. Y. Wu *et al*., Telomerase reverse transcriptase mediates EMT through NF-κB signaling in tongue squamous cell carcinoma. *Oncotarget* **8**, 85492–85503 (2017). doi: 10.18632/oncotarget.20888; pmid: 29156735
42. A. Hodorogea *et al*., Epithelial-Mesenchymal Transition in Skin Cancers: A Review. *Anal. Cell. Pathol. (Amst.)* **2019**, 3851576 (2019). doi: 10.1155/2019/3851576; pmid: 31934531
43. W. E. Damsky *et al*., β-catenin signaling controls metastasis in Braf-activated Pten-deficient melanomas. *Cancer Cell* **20**, 741–754 (2011). doi: 10.1016/j.ccr.2011.10.030; pmid: 22172720
44. G. S. Kinker *et al*., Pan-cancer single-cell RNA-seq identifies recurring programs of cellular heterogeneity. *Nat. Genet.* **52**, 1208–1218 (2020). doi: 10.1038/s41588-020-00726-6; pmid: 33128048
45. M. R. Webster, A. T. Weeraratna, A Wnt-er migration: The confusing role of β-catenin in melanoma metastasis. *Sci. Signal.* **6**, (2013). doi: 10.1126/scisignal.2004114; pmid: 2004114
46. W. E. Damsky, N. Theodosakis, M. Bosenberg, Melanoma metastasis: New concepts and evolving paradigms. *Oncogene* **33**, 2413–2422 (2014). doi: 10.1038/onc.2013.194; pmid: 23728340
47. A. Kaur, M. R. Webster, A. T. Weeraratna, In the Wnt-er of life: Wnt signalling in melanoma and ageing. *Br. J. Cancer* **115**, 1273–1279 (2016). doi: 10.1038/bjc.2016.332; pmid: 27764844
48. L. A. Garraway *et al*., Integrative genomic analyses identify MITF as a lineage survival oncogene amplified in malignant melanoma. *Nature* **436**, 117–122 (2005). doi: 10.1038/nature03664; pmid: 16001072
49. I. Tirosh *et al*., Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189–196 (2016). doi: 10.1126/science.aad0501; pmid: 27124452
50. M. Lauss, J. Nsengimana, J. Staaf, J. Newton-Bishop, G. Jönsson, Consensus of Melanoma Gene Expression Subtypes Converges on Biological Entities. *J. Invest. Dermatol.* **136**, 2502–2505 (2016). doi: 10.1016/j.jid.2016.05.119; pmid: 27345472
51. L. Jerby-Arnon *et al*., A Cancer Cell Program Promotes T Cell Exclusion and Resistance to Checkpoint Blockade. *Cell* **175**, 984–997.e24 (2018). doi: 10.1016/j.cell.2018.09.006; pmid: 30388455
52. S. Davidson *et al*., Single-Cell RNA Sequencing Reveals a Dynamic Stromal Niche That Supports Tumor Growth. *Cell Rep.* **31**, 107628 (2020). doi: 10.1016/j.celrep.2020.107628; pmid: 32433953
53. T. O. Jensen *et al*., Intratumoral neutrophils and plasmacytoid dendritic cells indicate poor prognosis and are associated with pSTAT3 expression in AJCC stage I/II melanoma. *Cancer* **118**, 2476–2485 (2012). doi: 10.1002/cncr.26511; pmid: 21953023
54. B. Johnston *et al*., Chronic inflammation upregulates chemokine receptors and induces neutrophil migration to monocyte chemoattractant protein-1. *J. Clin. Invest.* **103**, 1269–1276 (1999). doi: 10.1172/JCI5208; pmid: 10225970
55. R. Zilionis *et al*., Single-Cell Transcriptomics of Human and Mouse Lung Cancers Reveals Conserved Myeloid Populations across Individuals and Species. *Immunity* **50**, 1317–1334.e10 (2019). doi: 10.1016/j.immuni.2019.03.009; pmid: 30979687
56. A. M. De Silva, A. Gallardo, D. A. Fraser, Macrophage production and activity of innate immune proteins C1q, C1r, and C1s are modulated in response to molecular patterns. *J. Immunol.* **204**, 226 (2020).
57. D. Fraser, E. Melzer, A. Camacho, M. Gomez, Macrophage production of innate immune protein C1q is associated with M2 polarization (INM1P.434). *J. Immunol.* **194**, 56 (2015).
58. F.-C. Chou, H.-Y. Chen, C.-C. Kuo, H.-K. Sytwu, Role of Galectins in Tumors and in Clinical Immunotherapy, Role of Galectins in Tumors and in Clinical Immunotherapy, *Int. J. Mol. Sci.* **19**, 430 (2018). doi: 10.3390/ijms19020430; pmid: 29389859
59. K. Nazeri, A. Aminpour, M. Ebrahimi, "Two-stage convolutional neural network for breast cancer histology image classification" in *Lecture Notes in Computer Science*, B. ter Haar Romeny, A. Campilho, F. Karray, Eds. (Springer, 2018), vol. 10882, pp. 717–726.
60. J. N. Kather *et al*., Pan-cancer image-based detection of clinically actionable genetic alterations. *Nat. Can.* **1**, 789–799 (2020). doi: 10.1038/s43018-020-0087-6; pmid: 33763651
61. Y. Fu *et al*., Pan-cancer computational histopathology reveals mutations, tumor composition and prognosis. *Nat. Can.* **1**, 800–810 (2020). doi: 10.1038/s43018-020-0085-8; pmid: 35122049
62. O. Shalem *et al*., Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343**, 84–87 (2014). doi: 10.1126/science.1247005; pmid: 24336571
63. T. Wang, J. J. Wei, D. M. Sabatini, E. S. Lander, Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343**, 80–84 (2014). doi: 10.1126/science.1246981; pmid: 24336569
64. H. Koike-Yusa, Y. Li, E.-P. Tan, Mdel. C. Velasco-Herrera, K. Yusa, Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nat. Biotechnol.* **32**, 267–273 (2014). doi: 10.1038/nbt.2800; pmid: 24535568
65. L. A. Gilbert *et al*., Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* **159**, 647–661 (2014). doi: 10.1016/j.cell.2014.09.029; pmid: 25307932
66. A. Dixit *et al*., Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. *Cell* **167**, 1853–1866.e17 (2016). doi: 10.1016/j.cell.2016.11.038; pmid: 27984732
67. D. A. Jaitin *et al*., Dissecting Immune Circuits by Linking CRISPR-Pooled Screens with Single-Cell RNA-Seq. *Cell* **167**, 1883–1896.e15 (2016). doi: 10.1016/j.cell.2016.11.039; pmid: 27984734
68. C. Lee, J.-S. Kim, T. Waldman, Activated PI3K signaling as an endogenous inducer of p53 in human cancer. *Cell Cycle* **6**, 394–396 (2007). doi: 10.4161/cc.6.4.3810; pmid: 17329971
69. AACR Project GENIE Consortium, AACR Project GENIE: Powering Precision Medicine through an International Consortium. *Cancer Discov.* **7**, 818–831 (2017). doi: 10.1158/2159-8290.CD-17-0151; pmid: 28572459
70. D. B. Johnson *et al*., Acquired BRAF inhibitor resistance: A multicenter meta-analysis of the spectrum and frequencies,

clinical behaviour, and phenotypic associations of resistance mechanisms. *Eur. J. Cancer* **51**, 2792–2799 (2015). doi: 10.1016/j.ejca.2015.08.022; pmid: 26608120
71. E. E. Patton *et al*., Melanoma models for the next generation of therapies. *Cancer Cell* **39**, 610–631 (2021). doi: 10.1016/j.ccell.2021.01.011; pmid: 33545064
72. S. J. Gallagher *et al*., Beta-catenin inhibits melanocyte migration but induces melanoma metastasis. *Oncogene* **32**, 2230–2238 (2013). doi: 10.1038/onc.2012.229; pmid: 22665063
73. D. X. Nguyen, P. D. Bos, J. Massagué, Metastasis: From dissemination to organ-specific colonization. *Nat. Rev. Cancer* **9**, 274–284 (2009). doi: 10.1038/nrc2622; pmid: 19308067
74. A. Kawakami, D. E. Fisher, Bioinformatic Analysis of Gene Expression for Melanoma Treatment. *J. Invest. Dermatol.* **136**, 2342–2344 (2016). doi: 10.1016/j.jid.2016.09.011; pmid: 27884291
75. R. J. Sullivan, K. T. Flaherty, Resistance to BRAF-targeted therapy in melanoma. *Eur. J. Cancer* **49**, 1297–1304 (2013). doi: 10.1016/j.ejca.2012.11.019; pmid: 23290787
76. I. Arozarena, C. Wellbrock, Phenotype plasticity as enabler of melanoma progression and therapy resistance. *Nat. Rev. Cancer* **19**, 377–391 (2019). doi: 10.1038/s41568-019-0154-4; pmid: 31209265
77. H. Cirenajwis *et al*., Molecular stratification of metastatic melanoma using gene expression profiling: Prediction of survival outcome and benefit from molecular targeted therapy. *Oncotarget* **6**, 12297–12309 (2015). doi: 10.18632/oncotarget.3655; pmid: 25909218
78. L. Li, M. Fukunaga-Kalabis, M. Herlyn, The three-dimensional human skin reconstruct model: A tool to study normal skin and melanoma progression. *J. Vis. Exp.* **54**, e2937 (2011). doi: 10.3791/2937; pmid: 21847077
79. J. Lee *et al*., Hair-bearing human skin generated entirely from pluripotent stem cells. *Nature* **582**, 399–404 (2020). doi: 10.1038/s41586-020-2352-3; pmid: 32494013
80. R. L. Belote *et al*., Human melanocyte development and melanoma dedifferentiation at single-cell resolution. *Nat. Cell Biol.* **23**, 1035–1047 (2021). doi: 10.1038/s41556-021-00740-8; pmid: 34475532
81. A. M. Taylor *et al*., *Cancer Cell* **33**, 676–689.e3 (2018). doi: 10.1016/j.ccell.2018.03.007; pmid: 29622463
82. J. M. McFarland *et al*., Multiplexed single-cell transcriptional response profiling to define cancer vulnerabilities and therapeutic mechanism of action. *Nat. Commun.* **11**, 4296 (2020). doi: 10.1038/s41467-020-17440-w; pmid: 32855387
83. A. Liberzon *et al*., The Molecular Signatures Database, The Molecular Signatures Database hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015). doi: 10.1016/j.cels.2015.12.004; pmid: 26771021
84. E. Hodis *et al*., Code for Compositional Analysis for: Stepwise-edited, human melanoma models reveal mutations' effect on tumor and microenvironment. Zenodo (2022); doi: 10.5281/zenodo.6326008

ACKNOWLEDGMENTS

We thank D. Kotliar, I. R. Watson, O. Ursu, C-Z. Zhang, D. E. Fisher, M. L. Meyerson, D. M. Sabatini, and W. R. Sellers for helpful discussions; L. Gaffney and A. Hupalowska for assistance with figure design and graphics; I. Avraham-Davidi for assistance with tumor dissociation protocols; G. Getz, C. Stewart, and L. Lichtenstein for assistance with whole genome sequence analysis; and J. Pfiffer, D. Kelly, J. Barnett, M. Veneskey, and B. Buckley for laboratory operations. We thank B. E. Bernstein for support during the review process of this manuscript. We thank R. Platt and the F. Zhang laboratory for the kind gift of an AAV transfer plasmid backbone. We thank the MIT KI Swanson Biotechnology Center Histology Core and the Broad Genomics Platform for their services and expertise. The results published here are in part based upon data generated by the TCGA Research Network: https://www.cancer.gov/tcga. We acknowledge the American Association for Cancer Research (AACR) and its financial and material support in the development of the AACR Project GENIE registry, as well as members of the GENIE consortium for their commitment to data sharing. Funding: This work was supported by the Dr. Miriam and Sheldon G. Adelson Medical Research Foundation (L.A.G.), the Starr Cancer Consortium (L.A.G.), the Klarman Cell Observatory at the Broad Institute (A.R.), the Howard Hughes Medical Institute (A.R.), and Ludwig Cancer Research (A.R.). E.H. is grateful for support from the Harvard Herchel Smith Fellowship, the Harvard Landry Cancer Research Fellowship, the Paul and Daisy Soros Fellowship for New Americans, and an NIH Medical Scientist Training Program grant from NIGMS (PI: L. Walensky; T32GM007753). D.S. was funded by an Early Postdoc Mobility fellowship from the Swiss National Science Foundation (P2ZHP3_181475) and a Damon Runyon Fellowship (DRQ-03-20). Z.M. was supported by the NCI (R50-CA252138). L.A.G. and A.R. were investigators of the Howard Hughes Medical Institute while this work was conducted. Author contributions: Conceptualization: E.H., E.T.T., J.Y.H.K., T.B., L.A.G., A.R.; Investigation: E.H., E.T.T., J.Y.H.K., T.B., L.R.Z., S.P., J-C.H., L.B., T.M.D., D.P., D.D., L.T.N., D.S., Z.M., C.A.J., and A.H.; Supervision: O.R-R., M.C.M., L.A.G., and A.R.; Writing - original draft: E.H., E.T.T., J.Y.H.K., T.B., L.A.G., and A.R.; Writing - review and editing: all authors. Competing interests: E.H. is a consultant for and holds equity in Dyno Therapeutics and was a consultant for GV. T.B., J-C.H., D.P., O.R.R., L.A.G., and A.R. are employees of Genentech since 1 February 2021, 20 September 2021, 3 May 2021, 19 October 2020, 1 January 2019, and 1 August 2020, respectively. J.C.H. owns stock of F. Hoffmann-La Roche AG. D.S. is a consultant for Roche Glycart AG, since August 2021. A.R. and O.R.R. are coinventors on patent applications filed by the Broad Institute for inventions related to single-cell genomics. L.A.G. is an equity holder of Roche/Genentech and was a cofounder and equity holder at Foundation Medicine, Inc. and Tango Therapeutics. A.R. is an equity holder of Roche/Genentech and is a founder and equity holder of Celsius Therapeutics, an equity holder in Immunitas Therapeutics and until 31 July 2020, was an SAB member of Syros Pharmaceuticals, Neogene Therapeutics, Asimov, and ThermoFisher Scientific. E.H., L.A.G., and A.R. are named inventors on a patent application filed by the Broad Institute covering the work described in this manuscript (US Patent Application 16/631,916, National Phase of PCT/US2018/042737). Data and materials availability: Plasmids will be made available on AddGene (addgene.org/browse/article/28224839), subject to the Uniform Biological MTA. Immortalized cell lines will be made available upon request, to noncommercial entities, subject to the Uniform Biological MTA. Raw single-cell sequencing files are available in DUOS (DUOS-000136). Processed files are available and browsable on the Single Cell Portal (SCP1334). Code to reproduce the hierarchical Bayesian logistic mixed-effects model is available on Zenodo (*84*). All histopathological images and the associated machine-learning code are available on Image Data Resource (https://idr.openmicroscopy.org, accession number idr0135) and Code Ocean (https://codeocean.com/capsule/9921736/tree), respectively.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abi8175
Materials and Methods
Figs. S1 to S44
Tables S1 to S27
References (*85*–*115*)
MDAR Reproducibility Checklist
Data S1

1 April 2021; accepted 18 March 2022
10.1126/science.abi8175

RESEARCH ARTICLE SUMMARY

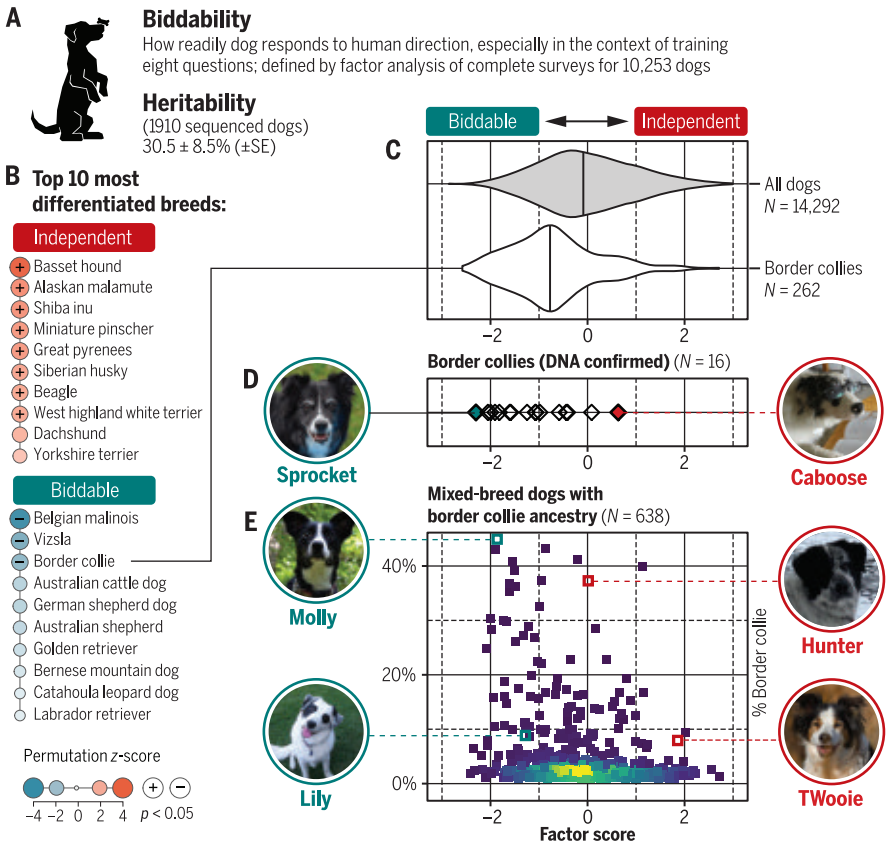
DOG GENOMICS

Ancestry-inclusive dog genomics challenges popular breed stereotypes

Kathleen Morrill*, Jessica Hekman, Xue Li, Jesse McClure, Brittney Logan, Linda Goodman, Mingshi Gao, Yinan Dong, Marjie Alonso, Elena Carmichael, Noah Snyder-Mackler, Jacob Alonso, Hyun Ji Noh, Jeremy Johnson, Michele Koltookian, Charlie Lieu, Kate Megquier, Ross Swofford, Jason Turner-Maier, Michelle E. White, Zhiping Weng, Andrés Colubri, Diane P. Genereux, Kathryn A. Lord, Elinor K. Karlsson*

INTRODUCTION: Before the 1800s, dogs were probably primarily selected for functional roles such as hunting, guarding, and herding. Modern dog breeds are a recent invention defined by conformation to a physical ideal and purity of lineage. Breeds are commonly ascribed temperaments and behavioral proclivities based on the purported function of the ancestral source population. By extension, the breed ancestry of individual dogs is assumed to be predictive of temperament and behavior. Through our community science project Darwin's Ark (darwinsark.org), we enrolled a diverse cohort of pet dogs to explore how genetics shapes complex behavioral traits in this exceptional natural model.

RATIONALE: Dogs are a natural system for investigating the genetics of complex traits. Millions of pet dogs live in human homes, sharing our environment, and receive sophisticated medical care. Behavioral disorders are treated with human psychiatric drugs, achieving similar response rates, and genetic studies suggest shared etiology with some human psychiatric conditions.



Effect of breed on behavior. (**A**) Biddability is among eight behavioral factors defined from surveys. SE, standard error. (**B**) Dogs in some breeds tend to score unusually high or low for this factor compared with dogs overall. (**C** and **D**) Border collies score lower on average for biddability (vertical line at median) but vary widely (C), including genetically confirmed border collies (D). (**E**) In mixed-breed dogs, border collie ancestry has a small genetic effect on biddability. [Photo credits: K. Wirka (Sprocket); M. Wisniewski (Caboose); B. Hoadley (Molly); M. Logsdon (Hunter); A. Macias (Lily); S. Staples (TWooie)]

We developed Darwin's Ark as an open data resource for collecting owner-reported phenotypes and genetic data and invited any dog owner to enroll their dog. We paired this with low-pass sequencing to capture nearly all common variation in this outbred population. Our inclusive approach achieved the large samples needed to investigate complex traits.

RESULTS: We surveyed owners of 18,385 dogs (49% purebred) and sequenced the DNA of 2155 dogs. Most behavioral traits are heritable [heritability (h^2) > 25%], but behavior only subtly differentiates breeds. Breed offers little predictive value for individuals, explaining just 9% of variation in behavior. For more heritable, more breed-differentiated traits, like biddability (responsiveness to direction and commands), knowing breed ancestry can make behavioral predictions somewhat more accurate (see the figure). For less heritable, less breed-differentiated traits, like agonistic threshold (how easily a dog is provoked by frightening or uncomfortable stimuli), breed is almost uninformative.

We used dogs of mixed breed ancestry to test the genetic effect of breed ancestry on behavior and compared that to survey responses from purebred dog owners. For some traits, like biddability and border collie ancestry, we confirm a genetic effect of breed that aligns with survey responses. For others, like human sociability and Labrador retriever ancestry, we found no significant effect.

Through genome-wide association, we found 11 regions that are significantly associated with behavior, including howling frequency and human sociability, and 136 suggestive regions. Regions associated with aesthetic traits are unusually differentiated in breeds, consistent with a history of selection, but those associated with behavior are not.

CONCLUSION: In our ancestrally diverse cohort, we show that behavioral characteristics ascribed to modern breeds are polygenic, environmentally influenced, and found, at varying prevalence, in all breeds. We propose that behaviors perceived as characteristic of modern breeds derive from thousands of years of polygenic adaptation that predates breed formation, with modern breeds distinguished primarily by aesthetic traits. By embracing the full diversity of dogs—including purebred dogs, mixed-breed dogs, purpose-bred working dogs, and village dogs—we can fully realize dogs' long-recognized potential as a natural model for genetic discovery. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: kathleen.morrill@umassmed.edu (K.M.); elinor.karlsson@umassmed.edu (E.K.K.)
Cite this article as K. Morrill *et al.*, *Science* **376**, eabk0639 (2022). DOI: 10.1126/science.abk0639

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abk0639

RESEARCH ARTICLE

DOG GENOMICS

Ancestry-inclusive dog genomics challenges popular breed stereotypes

Kathleen Morrill[1,2,3]*, Jessica Hekman[3], Xue Li[1,2,3], Jesse McClure[1], Brittney Logan[1,3], Linda Goodman[3,4], Mingshi Gao[1,2], Yinan Dong[1,3], Marjie Alonso[5,6], Elena Carmichael[1,3,7], Noah Snyder-Mackler[8,9,10], Jacob Alonso[3], Hyun Ji Noh[3], Jeremy Johnson[3], Michele Koltookian[3], Charlie Lieu[3,11], Kate Megquier[3], Ross Swofford[3], Jason Turner-Maier[3], Michelle E. White[1,3], Zhiping Weng[1], Andrés Colubri[1,3], Diane P. Genereux[3], Kathryn A. Lord[1,3], Elinor K. Karlsson[1,2,3,11,12]*

Behavioral genetics in dogs has focused on modern breeds, which are isolated subgroups with distinctive physical and, purportedly, behavioral characteristics. We interrogated breed stereotypes by surveying owners of 18,385 purebred and mixed-breed dogs and genotyping 2155 dogs. Most behavioral traits are heritable [heritability (h^2) > 25%], and admixture patterns in mixed-breed dogs reveal breed propensities. Breed explains just 9% of behavioral variation in individuals. Genome-wide association analyses identify 11 loci that are significantly associated with behavior, and characteristic breed behaviors exhibit genetic complexity. Behavioral loci are not unusually differentiated in breeds, but breed propensities align, albeit weakly, with ancestral function. We propose that behaviors perceived as characteristic of modern breeds derive from thousands of years of polygenic adaptation that predates breed formation, with modern breeds distinguished primarily by aesthetic traits.

Modern dog breeds are less than 160 years old (~50 to 80 generations), a blink in evolutionary history compared with the origin of dogs more than 10,000 years ago (*1*, *2*) (Fig. 1A). Prehistoric wolves likely adapted to use human refuse through changes in morphology, behavior, metabolism, and reproduction (*3*–*7*). Early humans may have given favored dogs increased access to limited resources, but there is little evidence of humans intentionally breeding dogs until 2000 years ago (*8*, *9*). By contrast, the modern dog breed, emphasizing conformation to a physical ideal and purity of lineage, is a Victorian invention (*10*). Before the 1800s, dogs were primarily selected for functional roles such as hunting, guarding, and herding (*11*)—heritable behaviors derived from the wolf predatory sequence (*12*). Modern breeds retain these component motor patterns, but their contexts, sequences, and thresholds vary (*12*, *13*). The extent to which ancient behavioral propensities persist in modern breeds, defined primarily by aesthetics and often disconnected from functional behavioral selection, is unclear.

Dogs with ancestry from a single modern breed (purebred dogs) predominate in genetic studies, which capitalize on their unusual population history and limited genetic diversity (*14*–*16*), but are a minority of all dogs (*3*, *17*). More than 80% of the nearly 1 billion dogs on Earth are free-living, free-breeding, and not under human control (e.g., village dogs) (*3*). Even in countries with large purebred populations, dogs with ancestry from more than one breed are common [~50% in the United States (*18*)]. Herein, we use the word "mutt" to describe a dog with ancestry from more than one breed, and potentially from nonbreed populations.

Modern breeds are commonly ascribed characteristic temperaments (e.g., bold, affectionate, or trainable), and behavioral proclivities on the basis of their purported ancestral function (e.g., herding or hunting) (*19*, *20*). By extension, the breed ancestry of an individual dog is assumed to be predictive of temperament and behavior (*21*), with dog DNA tests marketed as tools for learning about a dog's personality and training needs (*22*). Studies, however, found that within-breed behavioral variation approaches levels similar to the variation between breeds (*23*, *24*), suggesting that such predictions are error prone even in purebred dogs.

Behavioral traits in dogs are also a potentially powerful natural model for human neuropsychiatric disease. Pet dogs are regularly treated with human psychiatric drugs, including selective serotonin reuptake inhibitors, and have similar response rates, and genetic studies suggest shared etiology (*25*–*29*). Dog behavioral traits are polygenic, driven by many small effect loci and the environment (*30*, *31*). Given this genetic complexity, the success of survey-based phenotypes for mapping complex human diseases (*32*, *33*), and the availability of well-validated dog-owner surveys (*34*–*37*), a large-scale, survey-based study design is ideal for investigating the genetics of canine behavior.

Through our community science project Darwin's Ark (darwinsark.org), we enrolled a diverse cohort of pet dogs to explore the complicated, and sometimes unexpectedly weak, relationship between breed and behavior. We show that using ancestrally diverse dog cohorts enables more powerful studies of behavioral genetics in this notable natural model.

Results

Survey data

We developed Darwin's Ark as an open data resource for collecting owner-reported phenotypes and genetic data in dogs. Dog owners were asked to complete 12 short surveys (117 questions) on behavioral and physical traits (Fig. 1, B and C; figs. S1, S2A, and S3; and table S1). Darwin's Ark surveys, each with no more than 10 questions, are designed to be easy to complete. Owner survey responses are susceptible to rater bias, including the influence of breed stereotypes.

To ascertain size, we asked whether the dog is ankle-high, calf-high, knee-high, thigh-high, or hip-high rather than requiring owners to measure their dog (fig. S2A). This simple question proved effective when validated in three ways: (i) We measured dogs [$N = 38$ dogs; Pearson correlation coefficient ($R_{pearson}$) = 0.86; $p = 3 \times 10^{-12}$]; (ii) owners measured their dogs ($N = 337$ dogs; $R_{pearson} = 0.84$; $p = 6 \times 10^{-93}$); and (iii) we tested correlation with breed average height (*38*) in the subset of dogs that were purebred ($N = 2025$; $R_{pearson} = 0.85$; $p < 2.2 \times 10^{-16}$) (Fig. 1D, fig. S2, and data S1).

Owners answered, on average, 100 ± 34 (±SD) questions per dog; 70% of dogs have answers for more than 95% of questions (*22*) (Fig. 1, E and F). For the 48 questions drawn from the Dog Personality Questionnaire, between-question correlations matched published results (*22*, *37*) [Mantel's correlation coefficient (r) = 0.95, $p = 1 \times 10^{-7}$; fig. S1].

Overall, the Darwin's Ark cohort ($N = 18{,}385$; 85% from the United States) is broadly similar to the US dog population. Half (49.2%) are reported as purebred (*18*), and breed frequencies (Fig. 1G) correlate with US breed popularity (*39*) ($R_{pearson} = 0.88$; $p = 1.48 \times 10^{-32}$) (table S2).

To reduce the dimensionality of the survey data, we performed exploratory factor

[1]Bioinformatics and Integrative Biology, University of Massachusetts Chan Medical School, Worcester, MA 01655, USA. [2]Morningside Graduate School of Biomedical Sciences, University of Massachusetts Chan Medical School, Worcester, MA 01655, USA. [3]Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA. [4]Fauna Bio Inc., Emeryville, CA 94608, USA. [5]The International Association of Animal Behavior Consultants, Cranberry Township, PA 16066, USA. [6]IAABC Foundation, Cranberry Township, PA 16066, USA. [7]Rice University, Houston, TX 77005, USA. [8]Center for Evolution and Medicine, Arizona State University, Tempe, AZ 85251, USA. [9]School for Human Evolution and Social Change, Arizona State University, Tempe, AZ 85251, USA. [10]School of Life Sciences, Arizona State University, Tempe, AZ 85251, USA. [11]Darwin's Ark Foundation, Seattle, WA 98026, USA. [12]Program in Molecular Medicine, University of Massachusetts Chan Medical School, Worcester, MA 01655, USA.
*Corresponding author. Email: kathleen.morrill@umassmed.edu (K.M.); elinor.karlsson@umassmed.edu (E.K.K.)

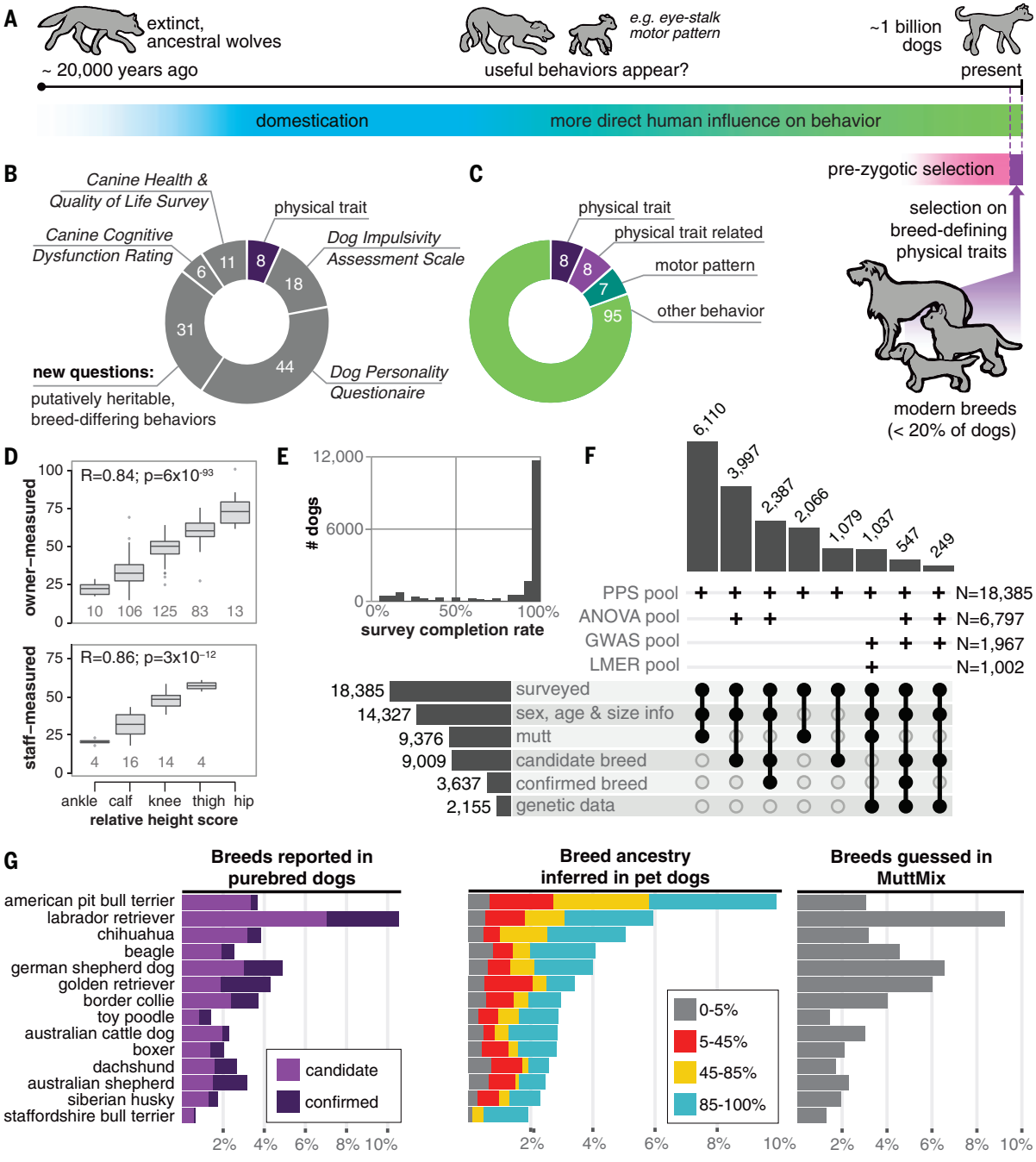


Fig. 1. The ancestry-inclusive Darwin's Ark project collected surveys and genetic data from a diverse cohort of dogs. (**A**) Selection on behavior in dogs predates modern breeds, which have existed for ~1% of dog history (*10*). (**B** and **C**) Surveys include (B) 79 published behavior questions (*34*–*37*) and 39 new questions about heritable, easy-to-identify traits that fall into (C) four categories that potentially vary in heritability, including physical traits (fig. S3). (**D**) Owner responses to the size question (fig. S2) were highly correlated with measured size in 375 dogs (Pearson correlation). Boxes and whiskers represent 25% quartile, 75% quartile, minimum, and maximum, with horizontal line at median. (**E**) Owners of most dogs answered >95% of questions. (**F**) Upset plot visualizing the dataset. Six horizontal bars show the number of dogs subsetted by information type or breed category ("subsets"). Each column is a possible intersection of subsets, with black circles indicating the inclusion of a subset and vertical bars showing the number of dogs intersecting. Four rows ("pools") represent the dogs used in four major analyses. Plus signs denote the inclusion of an intersection, with total number of dogs (*N*) on the right. (**G**) The frequency of breeds among purebred dogs in Darwin's Ark (left), compared with the 14 most common breed ancestries we detected in all dogs through global ancestry inference (middle). More popular breeds tended to be guessed more frequently by MuttMix participants (right).

analysis (*22*), defined eight factors that explain a cumulative 24.3% of variance in behavior (fig. S4 and tables S3 and S4), and scored 16,522 dogs. We named each factor for the facet of behavior captured and established standard terms for describing low- and high-scoring dogs (Fig. 2A). In subsequent analyses, we examined the influence of breed and genetics on each behavioral factor (Fig. 2B).

Genetic data

Nearly half of the Darwin's Ark cohort are mutts, an understudied population with uncharacterized genomic diversity. We whole-genome sequenced (mean coverage = 46×;

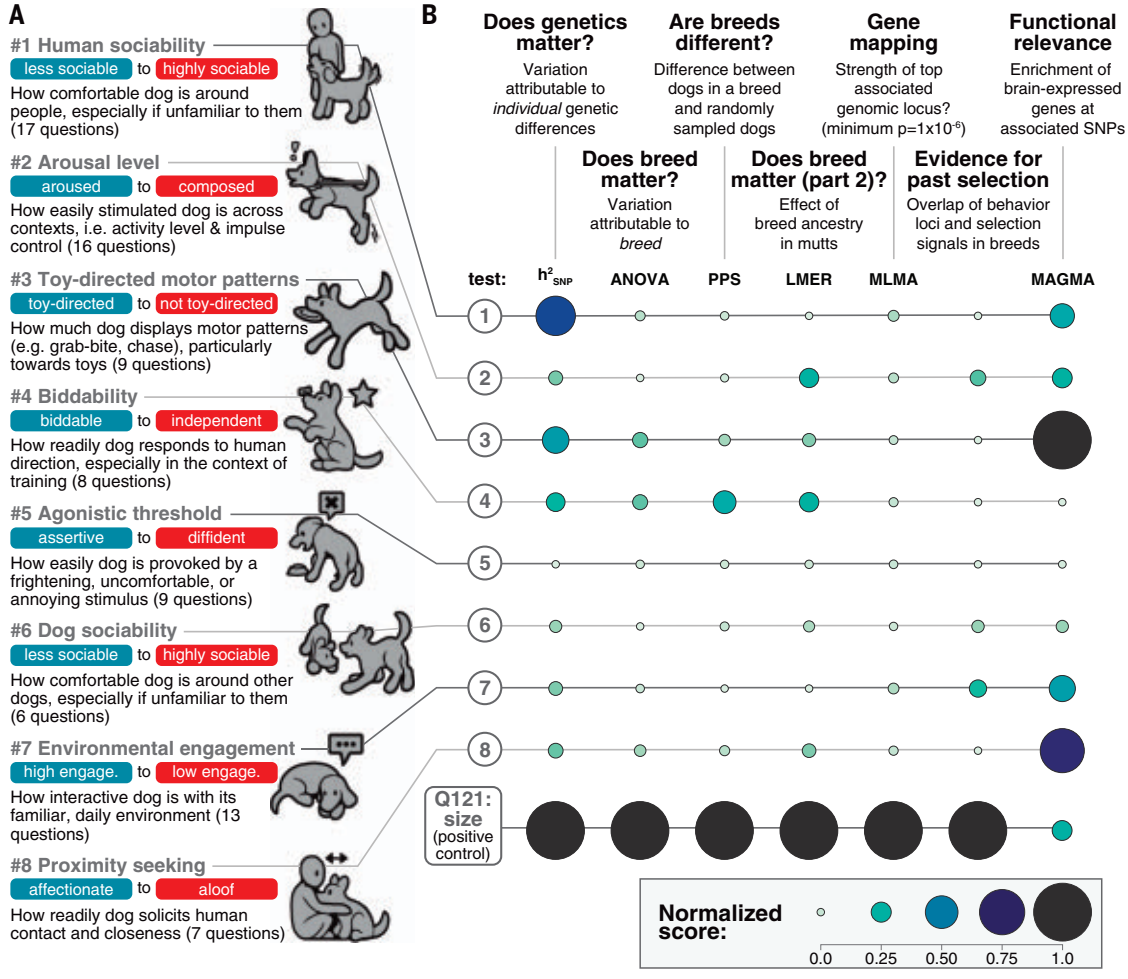

A
#1 Human sociability
less sociable to highly sociable
How comfortable dog is around people, especially if unfamiliar to them (17 questions)
#2 Arousal level
aroused to composed
How easily stimulated dog is across contexts, i.e. activity level & impulse control (16 questions)
#3 Toy-directed motor patterns
toy-directed to not toy-directed
How much dog displays motor patterns (e.g. grab-bite, chase), particularly towards toys (9 questions)
#4 Biddability
biddable to independent
How readily dog responds to human direction, especially in the context of training (8 questions)
#5 Agonistic threshold
assertive to diffident
How easily dog is provoked by a frightening, uncomfortable, or annoying stimulus (9 questions)
#6 Dog sociability
less sociable to highly sociable
How comfortable dog is around other dogs, especially if unfamiliar to them (6 questions)
#7 Environmental engagement
high engage. to low engage.
How interactive dog is with its familiar, daily environment (13 questions)
#8 Proximity seeking
affectionate to aloof
How readily dog solicits human contact and closeness (7 questions)
B
Does genetics matter?
Variation attributable to individual genetic differences
Are breeds different?
Difference between dogs in a breed and randomly sampled dogs
Gene mapping
Strength of top associated genomic locus? (minimum p=1x10-6)
Functional relevance
Enrichment of brain-expressed genes at associated SNPs
Does breed matter?
Variation attributable to breed
Does breed matter (part 2)?
Effect of breed ancestry in mutts
Evidence for past selection
Overlap of behavior loci and selection signals in breeds
test:
h2SNP
ANOVA
PPS
LMER
MLMA
MAGMA
1
2
3
4
5
6
7
8
Q121: size (positive control)
Normalized score:
0.0
0.25
0.50
0.75
1.0

Fig. 2. Behavioral traits do not define breeds the way aesthetic traits like size do. (**A**) Exploratory factor analysis bins behavioral questions into eight inferred factors, which correspond to latent behavioral propensities (blue, negative score; red, positive score). (**B**) In a series of seven analyses, we explored how behavior relates to breed in the context of size, which is a strongly breed-differentiated trait. For each column, circle size is proportional to the minimum-maximum normalized values of (i) LD-corrected h^2_{SNP}, (ii) effect size of breed in ANOVA (confirmed breed), (iii) standard deviation of PPS (candidate breeds), (iv) standard deviation of LMER *t* scores, (v) $-\log_{10}$ (minimum *p*) for MLMA, (vi) fraction of breeds with significant overlap ($p_{FDR} < 0.05$) between PBS and GWAS, and (vii) maximum MAGMA $\log_{10}(p)$ for 13 brain regions in GTEx (*85*).

range 30× to 67×) 27 pet dogs of unknown breed ancestry (the "Mendel's Mutts" cohort), including 26 from the United States (one originally from Mauritius and two from St. Kitts) and one from Ireland (data S2). We compared jointly called variant records in this cohort with published whole genomes for 530 purebred dogs from 128 breeds (*22*, *40*).

Sequencing mutts efficiently captures common variants in the dog population, including variation not detected by sequencing large numbers of purebred dogs. Sequencing an additional mutt yields nearly as many new variants as sequencing a purebred dog, even when each new purebred dog is from a different breed (Fig. 3A). In the 27 mutts, we discovered 11,974,853 biallelic single-nucleotide polymorphisms (SNPs) total, including 375,474 variants not found in the 128 breeds (530 dogs).

We also confirm that genetic variation private to a single breed is rare [38,097 ± 13,206 (±SD) SNPs per breed after excluding Tibetan mastiffs, a more distantly related lineage (*41*) with 651,551 private SNPs]. Breeds are not distinguished by a small number of "breed-defining" variants (table S5). After analyzing all 13 breeds with more than 10 sequenced representatives, we found just 332 SNPs (298 autosomal) exclusive to, and fixed in, a single breed (data S3) out of 16,702,091 SNPs total (0.002%). Tibetan mastiffs account for 142 SNPs (121 autosomal), with just 16 ± 31 (±SD) SNPs (15 ± 29 autosomal) in each of the nine other breeds. These variants are unlikely to affect phenotype. Annotation with SNPEff classifies 98.2% (326) as occurring at loci without obvious function.

Mendel's Mutts have shorter runs of homozygosity than purebred dogs (Fig. 3B) and linkage disequilibrium (LD) that decays more rapidly (fig. S5). Thus, genotyping arrays designed with sufficient marker density for purebred dog studies miss much of the genetic variation in mutts. Average squared correlation coefficient (r^2) between SNPs drops below 0.2 at 9.8 kb, which is slightly longer than in village dogs (6.2 kb) but 5- to 10-fold shorter than in breeds (fig. S5). Because of the short LD, the markers included on the Illumina Canine HD Beadchip (N = 171,882) and the Axiom Canine Genotyping Array Sets A and B (N = 1,011,992) tag only 19 and 53% of genomic variation in mutts, respectively, compared with 51 and 85% in breeds (Fig. 3C).

We adopted a low-pass sequencing and imputation approach (*42*–*46*), using a reference panel of 435 deeply sequenced dogs and other canids (data S4). We validated by resequencing 11 mutts with high-coverage whole-genome sequencing (WGS) at low coverage [1.0× ± 0.6× (±SD)]. We imputed, on average, 32,438,672 SNPs and 13,910,371 insertions and/or deletions (indels) per dog, or 19.8 ± 6.9 (±SD) variants per kilobase (~40× denser than the Axiom array), which was sufficient to tag nearly all the common variants (94% tagged by a marker within 100 kb and 87% within 1 kb) (Fig. 3C). Concordance between low-pass and 30× sequencing was 98.3 ± 0.7% (±SD) (N = 11 dogs; ~7.7 million common SNPs), which was slightly lower than that between the Axiom array and 30× (99.3 ± 0.1%; N = 10 dogs; 0.83 million SNPs) but better than that between imputed array calls and 30× (97.3 ± 0.3%; 7.6 million SNPs), including higher concordance for heterozygous genotypes (98.9 versus 98.3%) (data S5).

Our final genetic dataset comprises 1715 Darwin's Ark dogs sequenced at 0.6× ± 0.3× (±SD) coverage, each genotyped for 32,213,747 ± 141,060 (±SD) SNPs, and 440 dogs genotyped on the Axiom array and imputed using the same pipeline for 32,006,290 ± 157,307 (±SD) SNP genotypes, for a total of 2155 dogs (data S6).

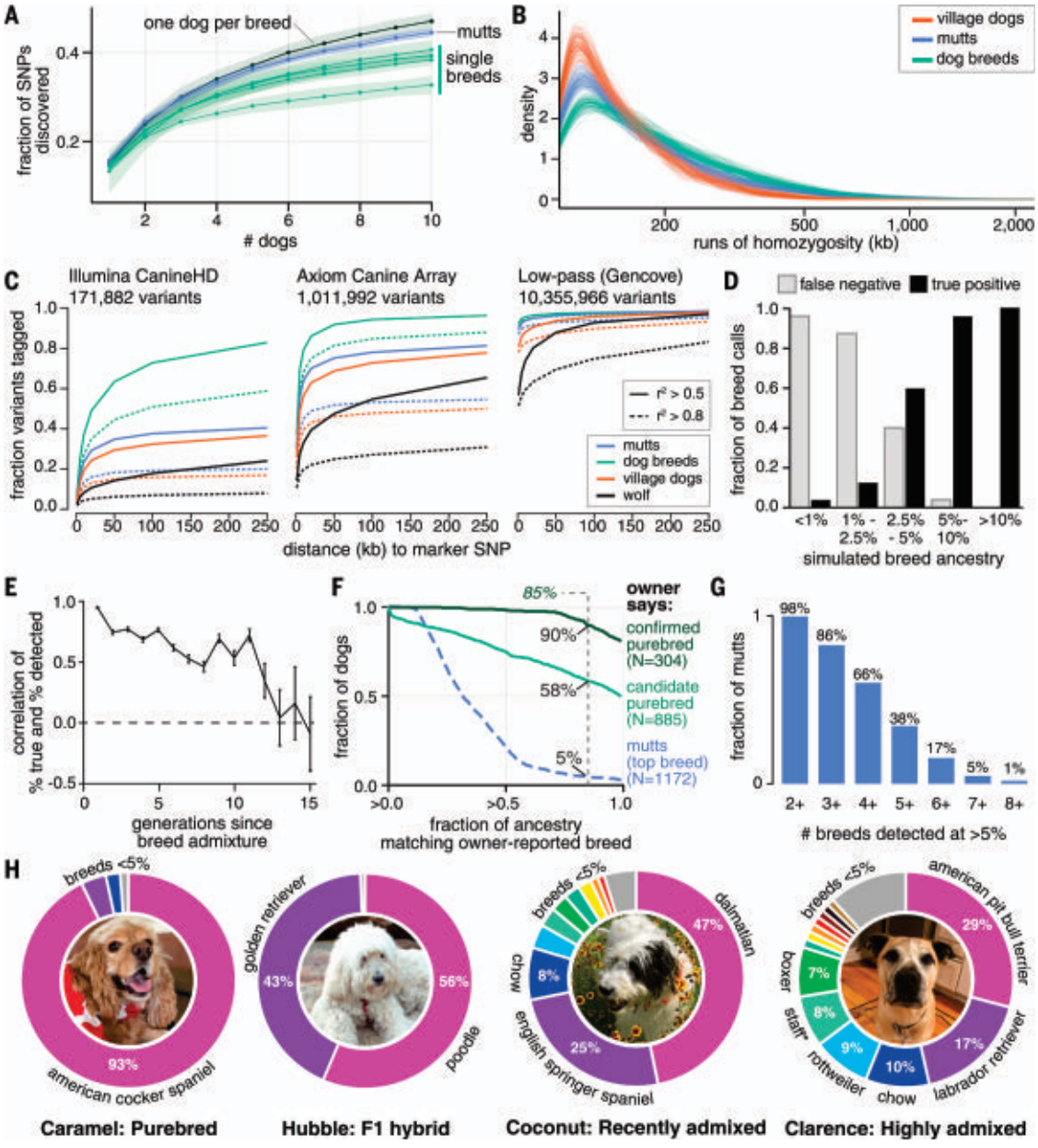


Fig. 3. Mutts have complex ancestry from many breeds. (**A**) Sequencing mutts discovers new variants at nearly the same rate as sequencing a series of purebreds of different breeds and faster than sequencing multiple individuals within any one breed (green; from top down: golden retriever, Labrador retriever, Yorkshire terrier, and Leonberger). Shaded regions indicate 95% confidence intervals from random reordering within each cohort. (**B**) Runs of homozygosity are shorter in mutts than in purebred dogs but are slightly longer than in outbred village dogs. (**C**) Fraction of variants tagged (out of a random sample of 20,000 autosomal SNPs) by a marker SNP is lower for the Illumina and Axiom genotyping arrays compared with low-pass sequencing with imputation, particularly in outbred populations like mutts. (**D** and **E**) Global breed ancestry inference pipeline (assessed using simulated breed admixture) (D) calls breeds comprising >5% ancestry accurately but misses lower-frequency breeds and (E) can discern admixture occurring up to ~12 generations ago (~24 to 36 years). Error bars in (E) represent standard deviation across 10 simulations of 100 admixed genomes. (**F**) For dogs categorized as confirmed purebred based on owner reports, breed calling assigns >85% ancestry (vertical dashed line) to the owner-reported breed in 90% of dogs. For candidate purebred dogs, 58% meet this criterion (4.4% had no detectable ancestry from the owner-reported breed). Just 5% of dogs categorized as mutts have >85% ancestry from their most common breed (blue dashed line). (**G**) Most mutts have ancestry detected (>5%) from more than three breeds (1205 dogs total). (**H**) Examples of breed calling in four dogs with different ancestry types: Caramel, a purebred dog, who has 93% of her ancestry assigned to her owner-reported breed; Hubble, an F_1 goldendoodle; Coconut, who has apparent mutt ancestry mixed with dalmatian; and Clarence, a mutt with <45% ancestry from any one breed (*Staffordshire bull terrier). [Photo credits: M. Movassagh (Caramel); J. Luban (Hubble); A. Pensarosa (Coconut); R. Skloot (Clarence)]

We selected dogs for sequencing on the basis of the enrollment date and survey completion rate (7.4% of dogs had sequencing funded by the owner's donation).

Breed ancestry assignment

In our genetic data, owner-reported breed is a reasonable proxy for predominant genetic ancestry. We developed a breed-calling pipeline using the software ADMIXTURE (*22*, *47*) to infer ancestry using a supervised analysis and a reference panel of 101 breeds (12 dogs per breed; 688,060 SNPs) collated from public and Darwin's Ark data (Fig. 3, D and E; figs. S6 and S7A; and table S2). Genetically inferred breed ancestry across dogs correlated well with the proportions of dogs registered to breeds in the American Kennel Club (93 breeds, $R_{pearson} = 0.74$; $p = 2.9 \times 10^{-17}$) (Fig. 1G).

The top breed that was called matched the owner-reported breed in 98.7% of dogs described as registered purebred ($N = 304$) and in 85.8% of all dogs for which owners report just a single breed ($N = 885$) (table S6). Dogs described as registered purebreds vary somewhat in the percent ancestry assigned to the owner-reported breed (potentially because the reference data are not representative of all diversity in the breed or because of shared ancestry between breeds). We empirically set the threshold for defining a dog as genetically "purebred" as when 85% of ancestry is inferred to come from a single breed because 90% of

dogs described as registered purebred by owners fell above this threshold (Fig. 3F).

We designated three classifications of breed ancestry: (i) "confirmed purebred dogs" were either described as registered purebred by the owner or confirmed by sequencing (3637 dogs), (ii) "candidate purebred dogs" included all confirmed purebred dogs and dogs with owner-reported ancestry from one breed (9009 dogs), and (iii) "mutts" were all other dogs (9376 dogs) (Fig. 1F). Genetically inferred ancestry superseded owner-reported breed when both were available, although discrepancies were rare (15 out of 556 candidate purebreds, 3 out of 323 confirmed purebred) and were primarily nominal variations on the same breed (e.g., Landseer versus Newfoundland). Extrapolating from the subset of dogs with genetic data, 89.7% of registered purebred and 58.2% of dogs with owner-reported ancestry from one breed would, if sequenced, have >85% ancestry called from their owner-reported breed (Fig. 3F). Because confirmed purebred dogs have a substantially higher percentage of their ancestry assigned to their owner-reported breed, in subsequent analyses, we prioritized the confirmed set or, when the larger and more diverse candidate set was useful, validated findings in the confirmed set.

Mutts are rarely (17%) mixes of just two breeds. Most (66%) carry >5% ancestry from four or more breeds (Fig. 3G). We find that 1071 dogs (70%) are highly admixed, carrying under 45% ancestry from any one breed (Fig. 3H). The most common breed ancestry (data S7) is American pit bull terrier (9.9%) followed by Labrador retriever (6.0%), Chihuahua (5.1%), beagle (4.1%), and German shepherd dog (4.0%) (Fig. 1G and fig. S7C), varying by geographic region (fig. S7D). Purebred dogs had higher coefficients of inbreeding, as estimated from the proportion of the genome in runs of homozygosity [F_{ROH} = 0.06 ± 0.04 (±SD); N = 633] than mutts (F_{ROH} = 0.02 ± 0.02; N = 1221) [for Student's t test p value ($p_{t\text{-test}}$) = 1.7×10^{-122}; t = 28.4, degrees of freedom (df) = 776.8] (fig. S7B).

Heritability of surveyed traits

Combining genetic and survey data for 1967 dogs, we found that genetic variation explains more than 25% of the variation in factor scores for human sociability, toy-directed motor patterns, and biddability (responsiveness to commands), as well as in responses to 38 of 83 (46%) behavioral questions and eight physical traits. We estimated SNP-based heritability (h^2_{SNP}) with standard errors using restricted maximum likelihood (*48*) and LD score correction (excluding 27 questions for which more than half of LD-stratified variance components were constrained) on genetic relationship matrices calculated for dogs with genetic data [8,518,951 autosomal SNPs with minor allele frequency (MAF) >2%] (data S8). Physical traits are exceptionally heritable, with five out of eight exceeding 85% heritability. Retrieving is the most heritable behavioral trait [52.5 ± 9.2% (±SE)], and human sociability is the most heritable factor (factor 1, 67.3 ± 13.0%). Behaviors related to intrinsic motor patterns and physical traits are more heritable than other behaviors (Fig. 4A). To assess whether these heritabilities are overestimated because of correlation between traits and breed ancestry, we recalculated them by incorporating the top 10 principal component eigenvectors of genetic variance (*49*, *50*). For the most part, we saw little change [median fold change of +0.02, 25% quartile = −0.049 and 75% quartile = +0.140] between estimates ($R_{pearson}$ = 0.97, p = 1.1×10^{-58}). Heritability decreases the most for biddability (factor 4, drops from 30.5 ± 8.5% to 20.0 ± 8.8%) and "circles before pooping" [question 64 (Q64), drops from 25.1 ± 8.1% to 8.0 ± 7.6%] (fig. S8).

Breed explains some behavior variance

In the owner surveys, breed explains a larger fraction of the variance in behavior phenotypes (110 questions and eight factors) than size, sex, or age, but the effect is relatively small (Fig. 4, B and C; fig. S9A; and data S9). In an analysis of variance (ANOVA) of confirmed purebred dogs representing 78 breeds, the breed effect, measured as generalized eta squared (ges) (*51*), averages 0.089 ± 0.039 (±SD) (range 0.034 to 0.253), correlates with heritability ($R_{pearson}$ = 0.89; p = 7.9×10^{-44}) (fig. S9B), and is about fivefold higher for the physical traits characteristic of breeds than for behavioral traits (fig. S9C). The same analysis using the less stringent "candidate purebred" breed definition is nearly perfectly correlated with the confirmed purebred analysis ($R_{pearson}$ = 0.99, p = 5.2×10^{-102}; N = 125), with ges values ~30% lower (mean ratio = 0.70 ± 0.11) (fig. S9D).

Age explains little of the variation [0.018 ± 0.035 (±SD)] overall, but for a subset of traits it exceeds 0.05, including two factors (arousal level and toy-directed motor patterns) and nine questions, which include five designed to assess aging-related traits (*36*) (Fig. 4C and fig. S9E). Sex has little effect (0.009 ± 0.044), except for "lifts leg to urinate" (Q66; ges = 0.48). Size has virtually no effect (6.6×10^{-4} ± 8.6×10^{-4}; range 2.5×10^{-7} to 0.006).

Breed is not a reliable predictor of individual behavior

For several factors, score distributions for individual breeds differ from the distribution of all dogs (fig. S10), with at least a few breeds over- or underrepresented in the highest-scoring quartile (fig. S11 and data S10). These distributions are based on owner survey data that may be influenced by breed stereotypes and other factors, and differences are not necessarily genetic in origin. For example, for human sociability (factor 1), an individual Labrador retriever (1.4-fold), golden retriever (1.6-fold), American pit bull terrier (1.4-fold), or Siberian husky (1.7-fold) was more likely to score in the highest quartile than a randomly selected dog, whereas a German shepherd dog (0.78-fold), Chihuahua (0.72-fold), or dachshund (0.56-fold) was less likely. Even so, in every breed represented by 25 or more dogs, the majority scored within one SD of the Darwin's Ark cohort mean (67.2 ± 7.5% within one SD and 95.4 ± 3.0% within two SD for confirmed purebred dogs). Behavioral factors show high variability within breeds, suggesting that although breed may affect the likelihood of a particular behavior to occur, breed alone is not, contrary to popular belief, informative enough to predict an individual's disposition.

We developed an interactive dashboard (https://darwinsark.org/muttomics) to illustrate the value offered by breed for predicting behavior in any individual dog. For example, the chance that an owner scores an individual dog in the highest quartile for human sociability increases from 22% for a mutt to 40% if that dog is a golden retriever (fig. S11). Users can select one or a combination of characteristics, and the site dynamically updates to show the frequency in 23 breeds and in mutts.

Measuring breed peculiarity

We developed a permutation-based approach to measure when dogs of a particular breed are described by owners as having behavioral characteristics that significantly differentiate them from other dogs. For each phenotype, we compared dogs within each breed to dogs sampled randomly from the full cohort, producing a "population peculiarity score" (PPS) (*22*). We tested both "confirmed" breeds (sample n = 50; up to six breeds) and "candidate" breeds (sample n = 25; up to 62 breeds) (Fig. 4, D and E; figs. S12 and S13A; and data S11).

Overall, breeds were only subtly differentiated on behavioral phenotypes. In the confirmed purebreds, only 5.1% (30/583) of breed-phenotype pairs were significantly differentiated for behavioral questions, compared with 41.5% (17/41 pairs) for physical traits. Scores for behavioral questions were not more correlated with each other than were scores for physical questions (table S7). Intrinsic motor patterns and physical traits tend to be slightly more breed differentiated (fig. S13B).

No behaviors are exclusive to any breed (fig. S14). Even in the breed with the lowest howling propensity, confirmed Labrador retrievers (Q17; N = 241; 78.4% never howl), 8% of owners report that their Labrador howls sometimes, often, or always. Although 90% (53/57) of confirmed greyhounds are reported to never bury their toys (Q29), owners described three dogs as frequent buriers.

We used the same permutation approach to measure how behavior changes as dogs age

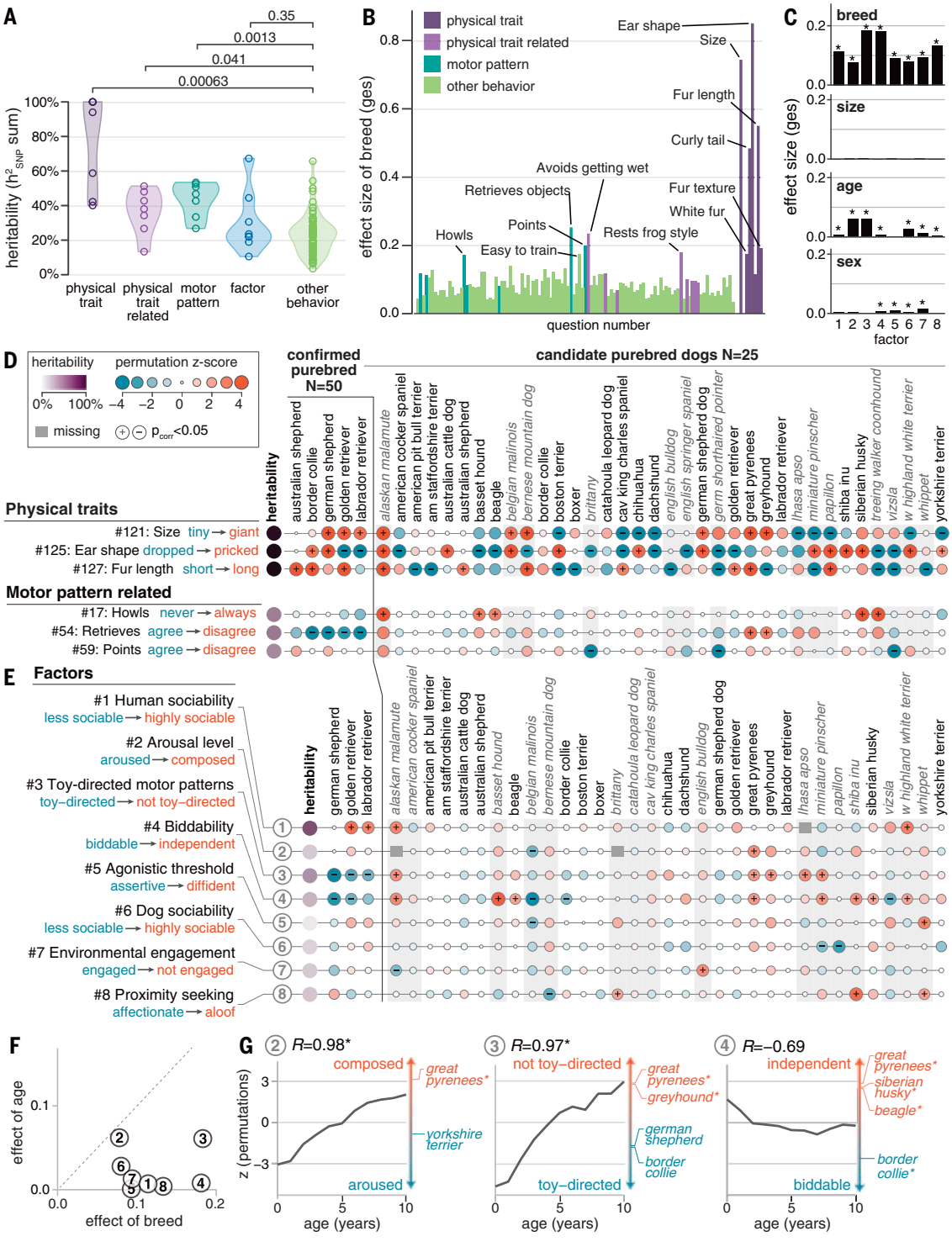


Fig. 4. Breed and age (and not size) have subtle effects on the behavior traits surveyed. (**A**) LD-corrected SNP-based heritability is much higher for physical traits, and somewhat higher for motor patterns, compared with other behavioral traits (significance measured using Student's *t* test; Benjamini-Hochberg (BH)–adjusted *p* value is shown). Shaded regions indicate the probability density. (**B** and **C**) ANOVA in confirmed purebred dogs shows that (B) the effect size of breed exceeds 15% for 6/7 physical traits and some behavioral questions (labeled bars) and that (C) for factors, breed explains more variation in scores than age and sex, and size has no significant effects. In (C), asterisks indicate statistical significance (BH-adjusted $p < 0.05$). (**D** and **E**) Permutations comparing breeds to randomly sampled dogs show that (D) many breeds are significantly differentiated on physical traits but that differentiation is much rarer for behavioral traits, including intrinsic motor patterns and (E) the eight behavioral factors. The selection of the most popular and/or most differentiated breeds is shown, with full results in fig. S12. Breeds in italics are represented by fewer than 50 individuals and skew toward more extreme *z*-scores (gray background; difference in mean score = 0.29; $p_{t\text{-test}} = 2 \times 10^{-71}$). (**F**) The effect of age, compared with the effect of breed, in the ANOVA shows that age explains nearly as much variance as breed for factor 2. (**G**) The PPS for arousal level and toy-directed motor patterns are significantly correlated with age, whereas biddability is more breed driven. Asterisks indicate significant results.

(fig. S13C). Most questions (72%) and factors (63%) are correlated with age [false discovery rate (FDR) *p* value (p_{FDR}) < 0.05] (figs. S15 and S16). Older dogs score nearly as composed in their arousal level (factor 2) as the most composed breed (Great Pyrenees), and puppies are far more toy-directed in their motor patterns (factor 3) than one of the most toy-directed breeds (German shepherd dog) (Fig. 4, F and G).

Testing breed stereotypes

PPSs aligned, to a limited degree, with behavioral stereotypes described in the breed standards (data S12). The American Kennel Club (AKC) describes each breed with a three-word

phrase (e.g., border collies are "affectionate, smart, energetic" and beagles are "friendly, curious, merry") (*52*) (table S2). Breeds described with particular words are not behaviorally distinct from other breeds; however, "charming" tends to describe breeds that are less toy directed (factor 3; $p_{FDR} = 0.039$) (fig. S17). Grouping breeds by their proposed historic working role, as captured by AKC show groups (*53*), finds that four out of six show groups are peculiar on at least one factor (Fig. 5A). Herding breeds are more toy directed, more biddable, more engaged, and more aloof, whereas toy breeds are more independent and less dog social. Sporting breeds are more toy directed, and working breeds are more dog social (Fig. 5B and fig. S18A).

We found more support for breed behavioral stereotypes when comparing the PPS results to quantitative rankings from the *Encyclopedia of Dog Breeds* for each breed on 10 behavioral characteristics (*19*) (fig. S17B). Nine of the 10 correlated significantly with PPS for at least one factor. Breeds ranked high on "ease of training" tended to be more biddable (factor 4) and more toy directed (factor 3). Breeds ranked as low on "energy level" scored as more composed, more dog social, and less environmentally engaged (factors 2, 6, and 7). "Watchdog ability" and "friendliness towards strangers" both correlated with human sociability (factor 1), but in opposite directions.

Overall, when comparing breeds to all pet dogs, breed differences based on owner reports align with some breed behavioral stereotypes, with one major caveat. Using survey data alone, we cannot distinguish environmental effects, including the effects of the stereotypes themselves (e.g., by influencing owner's perception of their dog's behavior), from genetic effects.

Human perception of breed in mutts

Half of the Darwin's Ark cohort were mutts, offering an opportunity to test whether breed stereotypes have a genetic etiology. In purebred dogs, cultural breed stereotypes affect the perception of a dog's behavior and thus may alter a dog's environment (*54*, *55*) or introduce rater bias into owner survey responses. If breed ancestry is not readily discernible in mutts, these nongenetic factors would be mitigated, allowing us to discern the genetic effects of breed from human perception and other environmental factors.

To measure how accurately breed can be discerned from physical characteristics in mutts, we ran a 2-month community science project, MuttMix (muttmix.org) that recruited 26,639 participants (Fig. 5C). For 30 mutts with complex genetic ancestry (i.e., no first-generation crosses) (*22*), we asked participants to guess their top three breeds. Between 13,662 and 14,160 participants submitted responses for each dog. They accurately identified, on average, 20.9 ± 20.4% (±SD) of each dog's breed ancestry (fig. S19A). Breeds comprising a smaller proportion of a dog's ancestry were especially challenging to identify ($R_{pearson} = 0.61$; $p = 3.3 \times 10^{-10}$) (fig. S19B). Thus, any effect of perceived ancestry on survey responses is likely to be substantially mitigated in mutts.

The physical characteristics associated with a breed, like short fur (American pit bull terrier), short legs (dachshund), or pricked ears (Chihuahua), influenced how participants guessed, but this is an error-prone approach (Fig. 5, D to F, and fig. S19C). Dogs with ancestry from more popular breeds had more correct guesses because participants tended to guess popular breeds more frequently (with the exception of the underguessed American pit bull terrier) (Fig. 1G and fig. S20). To control for this, we calculated how often we would expect to see each possible combination of breed guesses if breed guess rates matched the population frequencies in Darwin's Ark (table S2) and compared the observed rate of correct guesses to the expected rate. Participants guessed correctly more often than expected for 73% of dogs (Fig. 5C and fig. S21A). The number of correctly guessed breeds by each participant for each dog was slightly higher for self-described dog professionals [$N = 84{,}918$; mean = 0.93 (SD 0.74)] than nonprofessionals [$N = 333{,}614$; mean = 0.81 (SD 0.72)] (fig. S21B).

Effect of breed ancestry in mutts

We measured whether breed influences behavior through genetics by examining only mutts with <45% ancestry from any single breed. This analysis illustrates the power of including complex mixes in behavioral studies, because any influence of breed stereotypes is mitigated when true breed ancestry is not readily discernible from appearance. We built linear mixed-effects regression (LMER) models for all factors and questions (breed ancestries as fixed effects; age and pairwise genetic relatedness as random effects) (*22*) (Fig. 6, A to C; fig. S22; and data S13). The proportion of variance in factor scores captured by genetic breed ancestry (marginal R^2) averaged 9 ± 3% (±SD), suggesting a weak but discernible genetic effect of breed on disposition (fig. S23B and data S14). Breed ancestry explained, on average, 20 ± 12% (±SD) of variance for physical traits in mutts.

We validated the LMER method by confirming that the ancestry effects for physical traits matched the breed standards for physical appearance (*38*). For example, ancestry from nine breeds (six with long fur and three with short fur) had a significant effect on fur length in mutts, and for each breed, the direction of effect matches the breed standard. In total, we assessed 51 breed-trait pairs for four traits (size, white coat color, ear shape, and fur length). The direction of effect matched in 50 out of 51 (table S8). For the size question (Q121), the LMER score is strongly correlated with the breed average height ($R_{pearson} = 0.86$; $p = 2.3 \times 10^{-6}$; $N = 19$).

Correlation between the LMER results and the PPSs confirm that some behavioral differences in mutts derive from differences in breed ancestry ($N = 6333$; $R_{pearson} = 0.28$; $p = 1.8 \times 10^{-111}$) (fig. S24 and data S15). For example, mutts with more border collie ancestry tend to be more biddable (factor 4; LMER $t = -4.6$; $p_{FDR} = 0.0002$), consistent with survey data for confirmed border collies [PPS $z = -4.6$; corrected p value (p_{corr}) = 2×10^{-6}]. Similarly, mutts with more Labrador retriever ancestry tend not to avoid getting wet (Q60; LMER $t = 3.8$; $p_{FDR} = 0.003$), like many confirmed Labrador retrievers (PPS $z = 4.3$; $p_{corr} = 0.003$).

Discordance between the LMER results and breed differentiation measured by PPS may capture nongenetic influences on survey responses such as breed stereotypes. Owners of confirmed golden retrievers, for example, tend to disagree that their dog is fearful of unfamiliar people (Q46; PPS $z = 4.6$; $p_{corr} = 0.002$), which fits the breed stereotype that golden retrievers are friendly to strangers (*19*). In mutts, however, golden retriever ancestry had no effect on this question (LMER $t = 0.3$; $p_{FDR} = 0.88$), suggesting that the reported propensity may not be driven by genetics (fig. S23A). Similarly, whereas owners of confirmed Labrador retrievers tend to describe their dogs as more human social (factor 1; PPS $z = 3.3$; $p_{corr} = 0.006$), in line with the breed stereotype ("friendly, active, and outgoing"), in mutts, Labrador retriever ancestry has little effect (LMER $t = 0.4$; $p_{FDR} = 0.90$). Owners of confirmed border collies tend to score their dogs higher on "wants to play" (Q2; PPS $z = -3.6$; $p_{corr} = 0.04$); this is consistent with the stereotype that border collies are "affectionate, smart, and energetic" but discordant with the LMER results, which find no effect of border collie ancestry (LMER $t = 0.089$; $p_{FDR} = 0.97$)

By analyzing the effect of ancestry on behavior in mutts, we can anticipate the likely behavioral propensities of breeds that are not well represented in our survey data (table S9). For example, Saint Bernard ancestry correlates with being more affectionate (factor 8, LMER $t = -4.1$; $p_{FDR} = 0.002$) and Shar-Pei ancestry with being less toy directed (factor 3; LMER $t = 4.6$; $p_{FDR} = 0.0002$) (Fig. 6B). Ancestry from Chesapeake Bay retrievers increases with propensity to damage doors (Q40; LMER $t = 4.2$; $p = 0.001$) and escape from enclosures (Q35; LMER $t = 3.5$; $p_{FDR} = 0.02$).

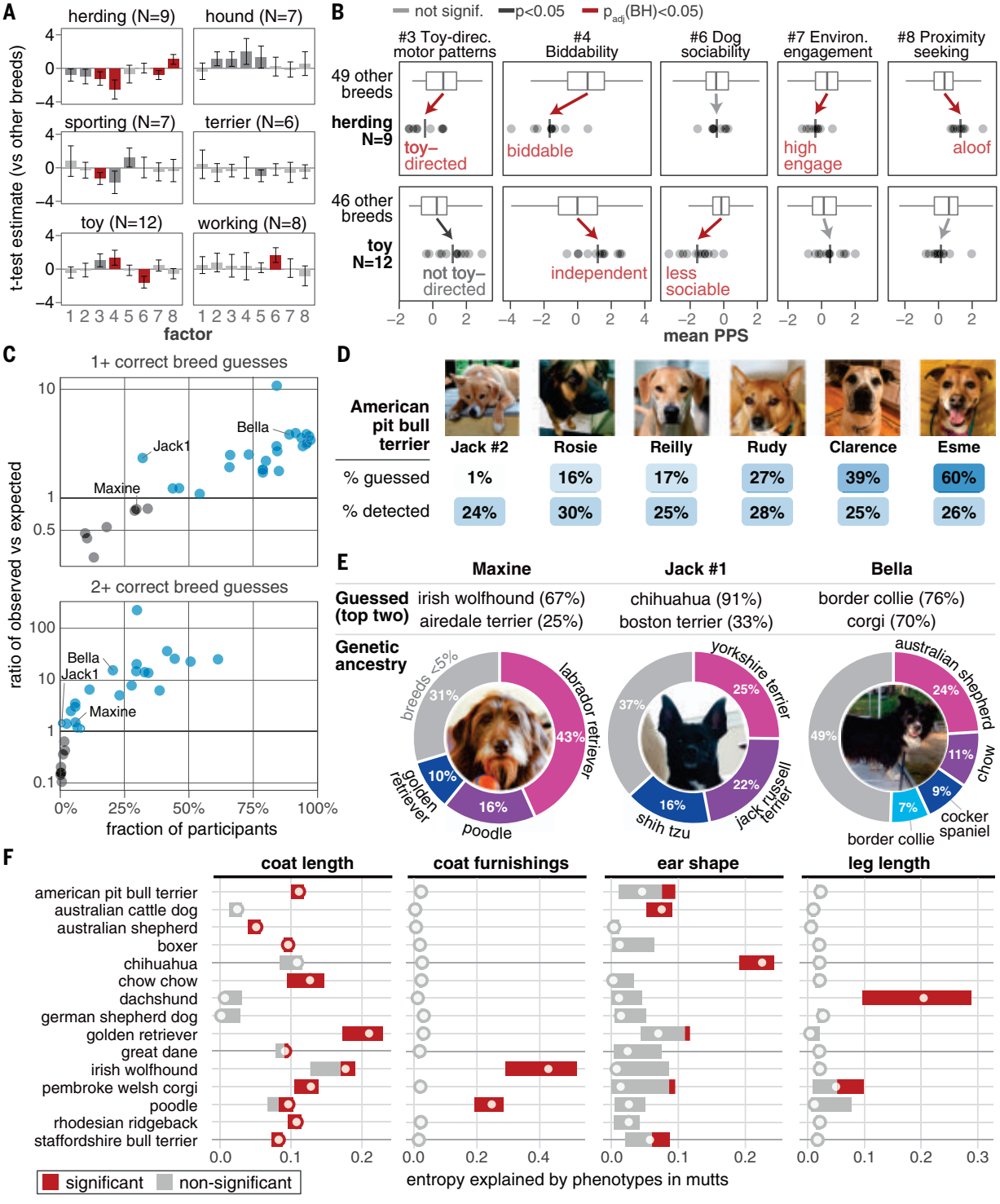


Fig. 5. Breed stereotypes can be assessed in mutts, where environmental effects (e.g., owner perception) are mitigated by the difficulty of accurately discerning the breed. (**A**) Breeds grouped by purported historic working roles are more differentiated (measured by mean PPS) on some factors than other breeds (58 breeds total), measured as a *t* statistic using Student's *t* test. Error bars represent 95% confidence intervals. (**B**) Toy and herding breeds illustrate groups with shifts in opposite directions (full results in fig. S18). Points are PPS for breed, vertical lines are mean PPS for group, boxes enclose the 25 to 75% quartiles, and horizontal lines extend from 1.5 times the interquartile range (IQR) below the 25% quartile to 1.5 times the IQR above the 75% quartile. Arrows indicate the direction of change in means, and words show the favored behavioral propensity. Herding breeds are Australian cattle dog, Australian shepherd, Belgian malinois, border collie, Catahoula leopard dog, collie, German shepherd dog, Pembroke Welsh corgi, and Shetland sheepdog. Toy breeds are bichon frise, Cavalier King Charles spaniel, Chihuahua, Havanese, Maltese, miniature pinscher, papillon, Pomeranian, pug, shih tzu, toy poodle, and Yorkshire terrier. (**C**) For each dog in the MuttMix survey, the ratio of observed:expected correct guesses for (i) one or more or (ii) two or more of the highest-content breed ancestries (blue indicates a ratio >1; open circles are not significant). (**D**) For six dogs that have a similar amount of genetic ancestry

detected from American pit bull terrier (~25 to 30%), participants guessed this breed at rates ranging from 1 to 60%. [Photo credits: J. O'Donnell (Jack); T. Fortier (Rosie); A. Phelps (Reilly); L. Moses (Rudy); R. Skloot (Clarence); M. Bishop (Esme)] (**E**) For three individual mutts, the most guessed breeds (top) differ from the genetically inferred breed ancestry (bottom). [Photo credits: E. Winchester (Maxine); R. Bacon (Jack); E. Stackpole (Bella)] (**F**) The dogs in (E) illustrate how a mutt's physical characteristics influence participant breed guesses. Points show entropy explained by traits using guesses for all mutts (*22*), and bars span values from a leave-one-out analysis (full results in fig. S18). For example, 67% of participants likely guessed Irish wolfhound for Maxine because of her coat furnishings.

Genome-wide association studies in all dogs

We first investigated breed-defining physical traits with known large-effect loci using a mixed linear model–based approach for genome-wide association (*56*) across 8,518,951 SNPs of >2% MAF. We controlled for population and family structure and cryptic relatedness in our complex cohort (600 purebred dogs, representing 88 breeds, and 1496 mutts) using a genetic relationship matrix in a mixed-model framework. None of the genome-wide association studies (GWASs) had unusual genomic inflation [mean inflation factor ($\lambda_{GC0.5}$) = 0.985 ± 0.016 (±SD); range 0.960 to 1.03; N = 14], suggesting that the mixed-model framework controls for confounding due to population structure and other factors (*57*).

We successfully replicated 17 published associations for physical traits other than size (table S10), including for genes *MITF* (*14*) with white spotting [$p = 2.89 \times 10^{-37}$; SNP effect size (b) = −0.78], *FGF5* and *RSPO2* (*58*) with coat length ($p = 5.46 \times 10^{-54}$; b = +0.37) and texture ($p = 6.35 \times 10^{-9}$; b = +0.11), *USH2A* (*59*) with roan and/or ticking ($p = 5.31 \times 10^{-16}$; b = +0.20), *RUNX3* (*60*) with pheomelanin intensity ($p = 4.11 \times 10^{-8}$; b = −0.20), and the β-defensin region (*61*–*63*) with brindle coat patterns ($p = 2.50 \times 10^{-107}$; b = +0.35) (fig. S25, D to I, and data S16).

For size, a quantitative trait, we replicated 10 previously published associations (*40*, *64*–*71*) (Fig. 6D, fig. S25A, and table S10) and found new associations to *SAR1B* [$p = 2.01 \times 10^{-8}$; b = +0.12; metabolic disorders (*72*, *73*)] and *ANAPC1* [$p = 4.11 \times 10^{-8}$; b = +0.15; short stature in Rothmund-Thomson syndrome (*74*)]. By comparing giant dogs (N = 55) and then tiny dogs (N = 55) to average dogs (N = 1841), we distinguished variants associated with gigantism (fig. S25B) and dwarfism (fig. S25C) specifically. The *FGF4* retrogene locus, previously associated with chondrodysplasia (*67*), is more strongly associated in the tiny GWAS ($p_{all} = 1.16 \times 10^{-26}$; $p_{tiny} = 1.15 \times 10^{-29}$), dwarfing all other loci.

The height associations were robust even in the absence of purebred dogs, suggesting that the all-mutt GWAS might offer equivalent power to one that includes purebred dogs. In dogs carrying less than 45% ancestry from any breed, a cohort with about half as many dogs (970 versus 1951), we identified all the major stature-associated loci ($R_{pearson}$ = 0.91; $p < 1 \times 10^{-8}$) as well as a new association in *LRIG3* ($p = 7.29 \times 10^{-10}$; b = −0.31), a gene involved in bone morphogenetic protein–mediated body-size regulation (*75*) (fig. S25J).

Genomic predictions for height based on the GWAS-identified variants perform well in both purebred dogs and mutts, reflecting the strong selection on size among dog breeds. For a random forest regression model built using 1730 dogs and 2733 size-associated SNPs ($p < 1 \times 10^{-5}$) (*22*), predictions carried a mean squared error of 0.3 (fig. S26) and 66% of predictions fell within ±0.5 units of the relative size score (fig. S2) ($R_{pearson}$ = 0.77, $p = 3.90 \times 10^{-305}$), with no drop in accuracy for predictions made on mutts [predicted and true values differed by 0.46 ± 0.35 (±SD) in purebreds versus 0.43 ± 0.36 (±SD) in mutts; $p_{t\text{-test}}$ = 0.08, t = 1.75, df = 832]. Randomly selected SNPs, by comparison, performed poorly, with a mean squared error of 0.5 (45% of predictions within ±0.5 units). Predictions for relative stature validated well against more precise measurements taken in person (N = 310 dogs; $R_{pearson}$ = 0.91, $p = 8.8 \times 10^{-117}$) (Fig. 6E and fig. S27).

Behavioral GWASs

Applying the same GWAS approach to the behavioral phenotypes identified 11 genome-wide significant ($p < 5 \times 10^{-8}$) (*76*) and 136 suggestive ($p < 1 \times 10^{-6}$) associations (data S16). As with physical traits, the behavioral GWAS had minimal genomic inflation [mean $\lambda_{GC0.5}$ = 0.995 ± 0.0087 (±SD); range 0.976 to 1.05; N = 118]. The associations for behavioral traits were weaker, consistent with a more complex genetic architecture. They have not yet been independently replicated. The most significant association, to "gets stuck behind objects" (Q36), mapped to a 380-kb region ($p = 8.36 \times 10^{-11}$; b = +0.54) (Fig. 6F and fig. S25K) containing *SNX29*, a gene associated with cognitive performance in human GWASs (*77*–*79*). "Dog howls" (Q17) mapped to an intergenic region ($p = 9.63 \times 10^{-11}$; b = +0.54) (Fig. 6G and fig. S25L) between *SLC38A11* and *SCN3A*, a voltage-gated sodium channel involved in the development of speech and language (*80*). The top association to a behavioral factor was for human sociability (factor 1), downstream of the gene *HACD1* ($p = 2.41 \times 10^{-8}$; b = −0.36) (Fig. 6H and fig. S25M), a regulator of long-term memory (*81*) that is also associated with centronuclear myopathies (*82*).

In our diverse cohort with dense genotyping data, associated regions are smaller than those discovered using intrabreed GWASs with sparser marker sets. We compared our behavior-associated regions to those found in an earlier study of a different complex trait (osteosarcoma) at the same linkage threshold ($r^2 > 0.8$). In the Darwin's Ark GWAS, associated regions extend to a median 5.6 kb (25 to 75% quartile = 2.0 to 14 kb, mean 16.8 kb) around suggestive ($p < 1 \times 10^{-6}$) behavioral loci and 5.7 kb at physical trait loci (1.4 to 22 kb, mean 26.2 kb). By contrast, intrabreed GWASs of osteosarcoma in three breeds with diverse population structures mapped at median ranges of 86 kb (25 to 75% quartile = 57 to 162 kb) in racing greyhounds, 54 kb (21 to 409 kb) in rottweilers, and 1 Mb (743 kb to 1.4 Mb) in Irish wolfhounds (*83*). This increased resolution may facilitate the search for causal variants. In the Darwin's Ark GWAS, we can distinguish a region on chromosome 10 that is associated with stature (76.2 kb at $r^2 > 0.8$; *HMGA2*; $p = 1.84 \times 10^{-24}$; b = −0.31) from one associated with ear shape (118.7 kb at $r^2 > 0.8$; *MSRB2*; $p = 6.02 \times 10^{-23}$; b = −0.33) that were previously linked in interbreed GWASs (*71*, *84*) (fig. S28).

The mixed-model association approach may not fully control for spurious association that arises when a trait differs between breeds. An association for "focused in distracting situations" (Q21) (chr32:4,512,005; $p = 1.0 \times 10^{-8}$; b = −0.22) (fig. S25N) mapped to a locus containing *FGF5*, a gene associated with long-coated breeds (*58*). This association was lost when we conditioned on the top coat length–associated SNP (chr32:4,509,367; p = 0.0001; b = −0.15) (fig. S29), which is linked to the top focus-associated SNP (r^2 = 0.33). The original association likely reflected the spurious difference in focus scores between dogs with shorter and longer coats ($p_{t\text{-test}}$ = 0.00023; 2012 dogs). Pleiotropy is unlikely because fur length explains almost no variation in focus scores (ANOVA ges = 0.0004; p = 0.35; N = 2456). Consistent with this, the focus association on chromosome 32 weakens (chr32:4,512,005; $p = 1.2 \times 10^{-6}$; b = −0.18) when we include the top 10 SNP-based principal components in the mixed model (fig. S30).

To assess whether spurious breed-trait correlations are a major confounder in our analyses, we reran all GWASs and included the top 10 principal components in the mixed model. Only 6% (3/48) of our genome-wide significant associations were lost ($p > 1 \times 10^{-6}$) (data S16). We also tested whether the top 75 regions associated with dog size (a highly breed-differentiated

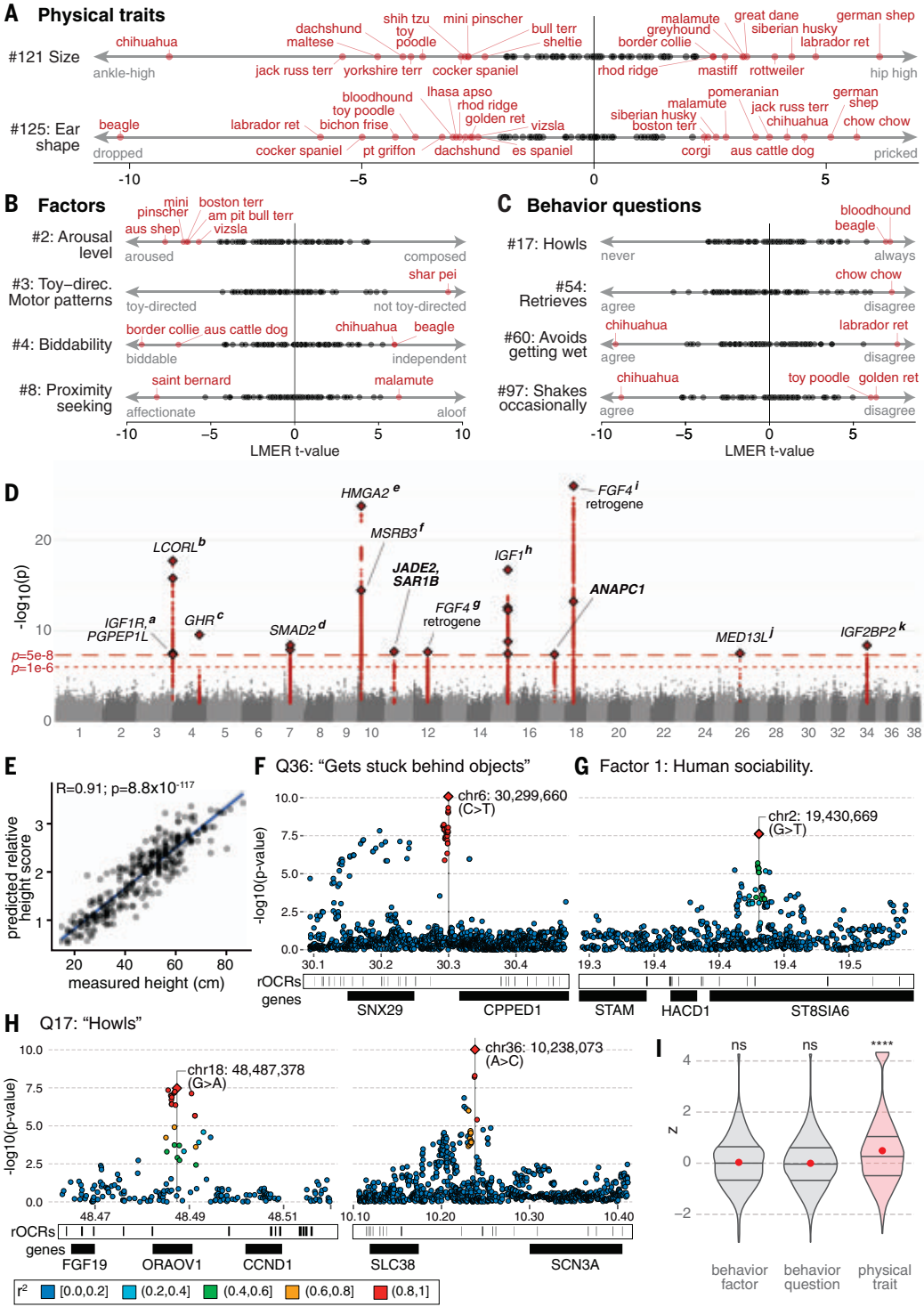


Fig. 6. Genetics of aesthetic and behavioral traits in dogs and influence of breed ancestry in mutts. (**A** to **C**) In highly admixed dogs with no breed ancestry over 45%, the fixed effects of breed ancestry on (A) physical traits, (B) behavioral factor scores, and (C) individual behavioral question scores are shown. (**D**) Manhattan plot for the GWAS of surveyed height (Q121) from 1951 dogs, including covariates for age and sex. Linkage blocks ($r^2 > 0.2$) associated ($p < 5 \times 10^{-8}$) with stature align with previous associations for body size in (a) *IGF1R* (*68*), (b) *LCORL* (*40*), (c) *GHR* (*69*), (d) *SMAD2* (*84*), (e) *HMGA2* (*69*) and the nearby (f) *MSRB3*, (g) a chromosome 12 retrogene insertion of *FGF4* (*70*), (h) *IGF1* (*64*–*66*), (i) another *FGF4* retrogene on chromosome 18 (*67*), (j) *MED13L* (*40*), and (k) *IGF2BP2* (*40*). Two previously unknown associations were found spanning *JADE2* and *SAR1B*, and in *ANAPC1*. (**E**) Random forest models based on size-associated SNPs ($p < 1 \times 10^{-5}$) accurately predict body size and correlate strongly ($N = 310$ dogs; $R_{pearson} = 0.91$, $p = 8.8 \times 10^{-117}$, $t = 37.451$, df = 308) with real measurements in those dogs. (**F** to **H**) Regional association plots for (F) scores on Q36 "gets stuck behind objects," (G) human sociability, and (H) frequency of howling from Q17. In addition to protein coding genes (black boxes), we also show representative open chromatin regions (rOCRs; narrow vertical lines). We annotated rOCRs genome-wide using ENCODE methods (*101*) applied to canine ATAC-seq (assay for transposase-accessible chromatin using sequencing) data from 14 tissues (*102*) and mammalian sequence constraint (*103*). (**I**) Breeds show high genetic differentiation, measured as the population branch statistic, overlapping physical trait associated loci compared with ~100,000 randomly permuted regions ($N = 1232$, mean $z = 0.49$, $p = 7.3 \times 10^{-33}$). Regions associated with behavioral factors ($N = 512$, mean $z = 0.03$, $p = 0.224$) and question scores ($N = 9317$, mean $z = 0.00$, $p = 0.603$) do not show such differentiation in breeds. Red circles indicate mean, with horizontal lines at the 25% quartile, median, and 75% quartile. The shaded area is the probability density, with significant differentiation in red. ns, not significant; ****$p < 0.0001$ (Student's t test).

trait) were enriched for SNPs associated in any of the 119 behavioral GWASs (fig. S31). Only one GWAS (Q66, "lifts leg to urinate") was significant [adjusted p value (p_{adj}) = 0.013]. Thus, although spurious associations due to aesthetic traits are a concern in multibreed GWASs, they are likely rare in the GWAS run on our diverse cohort.

Unattributed heritability

A large proportion of the genetic and environmental contributions to behavior remains undiscovered. The SNPs that we found to be associated with heritable ($h^2_{SNP} > 0.1$) behavioral traits account for a smaller proportion of overall heritability than do aesthetic trait associations (*22*), consistent with a complex genetic architecture. For the 14 physical traits,

53.0 ± 30.2% (±SD) of heritability is attributable to associated SNPs ($p < 1 \times 10^{-6}$), but for the eight behavioral factors and 73 questions, this drops to 21.0 ± 12.8% and 27.9 ± 20%, respectively. The six associated loci accounted for 42.7% of the genetic component of dog sociability (h^2_{SNP} = 14.8 ± 6.1%), whereas just 4.3% of highly heritable human sociability (h^2_{SNP} = 41.5 ± 9.1%) could be explained by its single associated region.

Brain-expressed genes enriched in behavior GWASs

Regions associated with dog behavioral phenotypes are enriched in brain-expressed genes. We cataloged the genes expressed in 38 tissue types, including 13 brain regions, using human GTEx data (*85*), an approach used previously in dogs (*86*). We also collated genes from curated lists for obsessive-compulsive disorder (OCD) (*87*), autism-spectrum disorders (*88*), and schizophrenia (*89*, *90*). Using MAGMA (*22*, *91*), we tested all GWASs for enrichment (data S17). Regions associated with toy-directed motor patterns (factor 3) had the strongest enrichments, which were for genes expressed in the hippocampus and in the basal ganglia of the nucleus accumbens, caudate, and putamen. Associations for "not keen on new situations" (Q84) were enriched for hypothalamus-expressed genes (fig. S30). Overall, enrichments in genes associated with neuropsychiatric conditions were weak, peaking for the enrichment of human OCD genes in Q84-associated regions (p = 0.0012; p_{adj} = 0.24).

Aesthetic selection predominates in breeds

Associations to physical traits, but not behaviors, tend to overlap signals of genetic differentiation in modern breeds, suggesting that aesthetics, and not behavior, has been the focus of selection. We tested for sites with excess differentiation in each breed with >12 dogs using the population branch statistic (PBS) test (*92*), using all dogs (N = 3802 to 3878) and wolves (N = 48) as the two outgroups across ~27.6 million SNPs from publicly available and our genetic data (data S4 and S6). Among the top 0.1% of breed-differentiated regions (26 ± 6 regions per breed), we validated genetic signals of selection reported at *EPAS1*, for hypoxia tolerance, in Tibetan mastiffs (*93*); at *CACNA1A*, unknown phenotype, in two sled dog breeds (*94*); at *ESR1*, unknown phenotype, in long-legged sighthounds (*40*); and at *ALX4*, a blue eye color gene, in Siberian huskies (*95*) (data S18).

We used permutation (*22*) to test whether PBS scores are unexpectedly high in regions associated with traits (data S19) and found that, whereas physical trait–associated regions are more differentiated, those associated with behavioral traits are not (mean z = 0.491 versus −0.001; $p_{t\text{-test}} = 4 \times 10^{-31}$) (Fig. 6I). Considering all moderately associated GWAS regions ($p < 1 \times 10^{-6}$), 25 of 65 (39%) physical trait loci are unusually differentiated, whereas only 38 of 515 (7%) behavioral trait loci are, and a subset of those are also connected to physical traits (data S20). Differentiation at physical trait loci is consistent with ongoing selection to meet strict morphometric standards in breeds (*38*), and the lack of overlap for most behavioral traits suggests weaker or absent selection.

The lack of differentiation at behavioral loci is not inconsistent with heritable behavioral differences in breeds, which may reflect genetic drift or selection that predates breed creation, neither of which the PBS test is designed to detect. To this point, neither of the two loci associated with howling are differentiated in either the Siberian huskies or beagles, even though ancestry from these breeds influences howling propensity.

Discussion

Behavioral traits are subtly differentiated in modern breeds (Fig. 2B). Furthermore, breed offers only modest value for predicting the behavior of individual dogs. For more heritable and more breed-differentiated traits, like biddability (factor 4), knowing breed ancestry can make behavioral predictions somewhat more accurate in purebred dogs. For less heritable, less breed-differentiated traits, like agonistic threshold (factor 5), which measures how easily a dog is provoked by frightening, uncomfortable, or annoying stimuli, breed is almost uninformative.

In our ancestrally diverse cohort, we show that behavioral characteristics ascribed to modern breeds are polygenic, environmentally influenced, and found, at varying prevalence, in all breeds. They likely naturally arose over millennia as dogs followed human migrations and adapted to new human technologies (*2*). The tight bottlenecks that established modern breeds captured ancient variation, at varying frequencies, with subsequent genetic drift or selection further shaping modern breeds (Fig. 3, A and B).

We found no evidence that the behavioral tendencies in breeds reflect intentional selection by breeders (Fig. 6I) but cannot exclude the possibility. Current datasets are too small to detect more subtle, recent directional selection, which requires hundreds of thousands of samples (*96*). In dogs, breed demographic history makes detecting selection particularly challenging (*1*, *97*).

Canine behavioral disorders are a proposed natural model for human neuropsychiatric diseases (*25*, *27*). Here, we show that large-scale behavior GWASs in dogs are tractable, identifying dozens of loci associated with behavioral traits in dogs. These associations explain a fraction of overall heritability, suggesting that still-larger sample sizes are needed. Our study design, combining owner-engagement with low-pass sequencing (*45*), makes this eminently achievable. We anticipate that this approach will be even more powerful once methods for accurately assigning local ancestry in individuals with >100 potential source populations (compared with two or three in human studies) are validated and incorporated into dog GWASs (*98*).

As dog studies grow in scale and complexity, it is crucial that we meet the standards of statistical rigor developed by the human genetics community and carefully account for confounding by artificial selection for aesthetic extremes in modern breeds (*99*), which can create misleading signals of association. One approach for studying behavior in dogs has been to compare breeds, rather than individuals, using breed-level behavioral phenotypes. The wide variability in behavior within breeds, and the potential for spurious correlations with breed-defining aesthetic traits, suggests that any discoveries made using this approach should be carefully validated using other methods.

To date, dog genetics has focused on modern breeds, which capture just a tiny fraction of global canine diversity. Although this made early genomic studies feasible (*14*), it limits discovery today (*100*). By embracing the full diversity of dogs, including purebred dogs, mutts, purpose-bred working dogs, and village dogs, we can fully realize dogs' long-recognized potential as a natural model for genetic discovery.

Materials and methods summary

Materials and methods described in full detail can be found in the supplementary materials (*22*).

Survey data collection

We collected consent, profile information, and surveys for 18,385 dogs enrolled by their owners via the Darwin's Ark platform (https://darwinsark.org) on or before 15 November 2019. Profile information included the dog's approximate birth date, sex and sterilization status, suspected or known breed(s), purebred registration, and/or photograph. We collected 12 surveys, including 11 about behavior (10 questions each) and one about physical characteristics (eight questions), for a total of 118 survey items (table S1). All responses to survey questions were time stamped, and ages at the time of survey were calculated relative to reported birth date (*22*).

The 110 behavioral questions all used a five-point Likert scale: (i) 81 questions had options of strongly agree, agree, neither agree nor disagree, disagree, or strongly disagree; and (ii) 29 had options of never, rarely, sometimes, often, or always. We sourced 79 behavioral

questions from published and validated surveys: (i) Dog Personality Questionnaire (DPQ/DPQL; 45 questions) (*37*); (ii) Canine Health-related Quality of Life Survey (CHQLS; 11 questions) (*36*); (iii) Dog Impulsivity Assessment Scale (DIAS; 18 questions, including one also in DPQ) (*34*); and (iv) Canine Cognitive Dysfunction Rating scale (CCDR; six questions) (*35*). We validated the performance of behavioral surveys using a Mantel's test on the inter-item correlation distance ($d = 1 - |r|$) matrices between published data for 48 DPQ items ($N = 2556$ dogs) and our data. We included 31 new behavior questions developed with input from canine behavior professionals in the International Association of Animal Behavior Consultants. The physical characteristics survey used a variety of response types (table S1). Answers of "I'm not sure," "I don't know," "not sure," and "surgically cropped ears" (Q125) were excluded.

Dog size was measured through Q121: "When DOG is standing next to someone of average height, how high are HIS shoulders?" This question was validated in three ways (fig. S2 and data S1): (i) owners were provided with a measuring tape by mail and instructed to measure the height from their dog's shoulder to the ground using the provided measuring tape (337 dogs); (ii) dogs were measured (height to withers) by professionals during the 2017 Somerville Dog Festival in Somerville, MA (38 dogs); and (iii) owner-reported size was compared with average breed height (2025 purebred dogs).

We performed exploratory factor analysis on the behavioral surveys (10,253 dogs with responses for all 110 questions) and extracted the optimal number of factors as estimated by the Horn's parallel analysis and optimal coordinates heuristic methods (20 factors; table S3). A varimax orthogonal rotation was applied to generate a structure matrix with factor loadings for each item, and items with low pattern or structure loadings (less than ±0.3) were removed. We generated factor scores for 6269 additional dogs with responses to >80% of questions by populating missing responses through random sampling. The dog's age for each factor is the mean age for all responses to included questions.

Sample collection

Animal study protocols for saliva and blood collection from dogs were approved by the UMass Chan Medical School Institutional Animal Care and Use Committee (IACUC) (no. A-2520-18). We sent or gave owners saliva collection kits (DNA Genotek PG-100 saliva swabs) for sampling. For a subset of dogs, owners provided blood collected by their veterinarian. We selected dogs for sequencing primarily based on survey completeness and enrollment date. Of 1715 samples submitted for low-coverage DNA sequencing, 159 samples (7.4% of 2155 dogs in the genetic dataset) were funded by owner donations to the Darwin's Ark Foundation, a 501(c)(3) nonprofit organization (82-3942341).

High-coverage genome sequencing and analysis

We performed high-coverage [45× ± 10× (±SD)] WGS on samples from 27 putatively mixed-breed dogs (the Mendel's Mutts cohort) (data S2). For the initial 22 mutts sequenced, we performed joint variant calling with publicly available data for 620 other dogs and 34 canids (data S4) using the Genome Analysis Toolkit (GATK3) (*22*) on the CanFam3.1 reference assembly. The final variant call file contained 34,191,821 SNPs and 11,943,064 indels. For the five mutts sequenced later, genotypes were called for the same set of variants using GATK3 HaplotypeCaller.

We compared cumulative variant discovery using purebred versus mutt genomes using chromosome 13 as a random proxy for the whole genome. We tested six cohorts: one dog sampled at random per breed ($N = 128$ possible dogs), Mendel's Mutts ($N = 27$ dogs), and the four breeds with >27 individuals sequenced (*22*). We computed the cumulative distribution of the fraction of 619,031 variants discovered using 557 purebred dogs versus using 10 dogs randomly chosen and ordered within each cohort and computed the 95% confidence interval using random reordering within each cohort.

We compared the lengths of detected runs of homozygosity (ROH) in mutts, dog breeds, and village dog genomes across biallelic SNPs using PLINK v1.90b6.21 with a minimum length of 100 kb and 100 SNPs, with at least 1 kb per SNP (*22*). We then randomly sampled $n = 464$ runs (the mean number of ROH detected per mutt) from the pool of ROH detected in mutts, purebred dogs, and village dogs, resampling $N = 100$ times.

We measured LD in mixed-breed dogs (Mendel's Mutts), breeds (golden retriever, Labrador retriever, Leonberger, and Yorkshire terrier), village dogs, and wolves by randomly sampling 25 dogs from each cohort and, for 20,000 randomly sampled biallelic SNPs, measuring r^2 to all SNPs within 100 kb. We assessed tagging of genetic variation using genotyping arrays by measuring r^2 between the same set of random SNPs and the subset of SNPs on the array (171,882 for the Illumina HD Canine Genotyping Array and 1,011,992 for the Axiom Canine Genotyping Array Sets A and B).

Low-coverage sequencing and imputation

We piloted a low-pass sequencing and imputation approach (*42*–*46*) using a panel of reference haplotypes from high-coverage whole-genome sequences. Autosomal variant calls were inferred directly from sequencing reads through Gencove loimpute software (*46*) and a panel of reference haplotypes from publicly available WGS data [mean coverage 22.9× (SD 14.2×)] for 435 canids (data S4). The imputation process generated unfiltered genotypes for 32,438,672 SNPs and 13,910,371 indels with imputation genotype probability (GP) scores per genotype per dog. We validated performance by comparing low-pass sequencing and imputation [1.0× ± 0.6× (±SD)] to array data (Axiom array) and high-coverage WGS data for 11 mutts with high-coverage WGS at low coverage. We also performed down-sampling of high-coverage WGS and subsequent imputation by the same method.

We combined low-pass sequencing data for 1715 dogs [0.6× ± 0.3× (±SD)] with data for 440 dogs genotyped on the Axiom array and imputed using the same haplotype reference panel (excluding genotypes of GP < 0.7). After merging, we performed additional quality control based on MAF, call rate, and Hardy-Weinberg equilibrium and validated owner-reported sex (*22*). The final dataset included 8,518,951 biallelic, autosomal SNPs and 2155 dogs at a genotyping rate of 97.5% (1084 males and 1071 females).

Breed ancestry assignment

We assembled a reference panel of 101 of the most common dog breeds in the United States (table S2) using high-coverage WGS for 380 dogs of 74 breeds (data S4), low-coverage WGS for 115 dogs of 54 breeds, Axiom genotyping array data for 109 dogs of 43 breeds, and Illumina CanineHD arrays for 883 dogs of 90 breeds (*22*). For each breed, we selected 12 dogs for inclusion, prioritizing high-density raw data and genetic diversity within breeds. We imputed genotypes for low-density data using the 435-canid panel of reference haplotypes. We retained SNPs genotyped in more than 80% of dogs and at a MAF of at least 5%. Among ancestry-informative SNPs of Hudson's estimator of fixation index (F_{ST}) > 0.15 between breeds, we selected a dense set of 2,468,442 markers ($r^2 > 0.9$ within 5 kb) for admixture simulations and a sparser set of 688,060 markers ($r^2 > 0.5$ within 50 kb) for ancestry inference.

We used a Monte Carlo approach to generate simulated admixed genomes of known ancestral haplotypes and then compared the breed ancestry composition with ancestry inferred using ADMIXTURE (*22*). We simulated admixed individuals through $N = 15$ generations of admixtures with the following procedure: $N + 1$ random individuals from different breeds were selected to contribute to the admixture. With each iteration, recombination was simulated to incorporate a new individual. Recombination was treated as a Poisson

event that occurred, on average, once every Morgan. Simulations ran on 10 independently drawn datasets of six dogs per reference breed to create 1000 admixed individuals of known ancestry. We inferred global ancestry for simulated individuals using the supervised mode of ADMIXTURE (random seed = 43) and the reference genotypes from six dogs reserved from each breed.

We then performed supervised admixture analysis of the Darwin's Ark genetic cohort. Genotype data from all query dogs was merged with all reference-breed data and filtered for SNPs in the global breed ancestry panel. Global ancestry from the 101 reference breeds were inferred using the supervised mode of ADMIXTURE (random seed = 43) that was supplied with reference population assignments. Population weights less than 1% were discarded from individual ancestry results.

We combined breed ancestry assignments with survey data for dogs without genetic data to define three breed sets as decribed in the results: confirmed purebred dogs, candidate purebred dogs, and mutts.

Heritability analysis

We estimated the SNP-based heritability (h^2_{SNP}) of surveyed traits using restricted maximum likelihood (REML) analysis implemented in the genome-wide complex trait analysis (GCTA, version 1.92.3 beta 3) software tool (*56*). We calculated LD scores in 250-kb regions using a block size of 10,000 kb with an overlap of 5000 kb between blocks. We generated a genetic relationship matrix (GRM) for the genetic cohort of 2155 dogs, as well as multiple GRMs calculated from SNPs stratified into LD score quartiles (*22*). The four LD-stratified GRMs were used to run REML analysis (GREML-LDMS) and estimate h^2_{SNP} with standard errors (data S8).

Population peculiarity scoring

We applied a custom permutation-based analysis (*22*) to test whether groups of dogs defined by breed or age differed significantly in survey responses from randomly sampled groups on any survey item or factor. We included all dogs with any survey responses. For each permutation and a given sample size N (table S14), we calculated the mean (the observed test statistic) for each normalized survey response or factor score for N dogs sampled from among dogs of each grouping. For each permutation, we also calculated the mean for a random sampling of size N from the full dataset (the permuted test statistics). We counted how often the observed test statistics for each population were higher than the permuted test statistics. We ran a total of 500,000 permutations. To obtain the PPSs, we calculated the one-tailed empirical p values and generated z-scores matching the survey directionality. We also calculated the two-tailed p values corrected for multiple testing by a maxT procedure that preserves the correlational structure between survey items (*22*).

Ancestry perception survey

We designed the web-based MuttMix survey (muttmix.org) to assess perceptions of breed ancestry in mixed-breed dogs by nonowner observers. Participants self-identified as either general public or dog professional (yes or no to "Do you work with dogs professionally and/or are you a breeder?"). The survey consisted of 30 mixed-breed dogs with ancestry assignments and one purebred dog. Owners provided front and side photographs and a short video. Owners reported the dog's relative size (fig. S2F) and other physical descriptors. The images and information that were provided were shared with participants, who were asked to guess, for each dog, the three breeds detected in largest proportion (*22*). The survey launched on 16 April 2018 and closed on 16 June 2018, and responses were collected from 26,639 people over a 2-month period.

We compared breed guesses to genetically inferred breed ancestry (*22*). Any breed call below 5% was removed and only breeds offered as survey options were examined. To calculate the average total percentage of ancestry guessed correctly, we first calculated the percentage guessed correctly by each user for each dog by summing the percent genetic ancestry attributed to their top three breed guesses. To assess the accuracy of user guesses of breed ancestry, we first counted the number of breed guesses for a given dog that were among the top two or three breeds that were genetically detected.

We measured how specific physical attributes affected participants' breed choices using entropy analysis (*22*). For each breed option, we calculated how well mutts' phenotypes, defined binarily for each of eight different traits (height, leg length, ear type, coat type, coat length, coat furnishings, white spotting, and pigmentation), distinguished between participant guesses of presence versus absence of each ancestry. We applied a leave-one-out analysis, omitting guesses for each mutt in series, to assess the impact of guesses for each mutt on entropy reduction. To calculate significance, we randomized trait assignments across mutts and then asked whether entropy reductions from true traits were greater than those randomly assigned.

We calculated how often we expected to see each possible combination of breed guesses by chance, assuming the guess rate for each breed to be the overall frequency of that breed (*22*) (table S2). We then calculated the observed rate of guesses with 1+, 2+, and 3 breeds correct for each dog and then calculated the ratio of the observed-to-expected rate.

LMER models

To measure the relationship of genetic breed ancestry to physical and behavioral phenotypes, we constructed LMER models using all dogs with <45% ancestry from any single breed (1002 dogs total). We treated normalized question and factor scores as independent variables, breed ancestry as fixed effects, and age as random effect. For each survey item and factor, we built a model with REML to obtain unbiased estimates, standard deviations, and Wald statistics (t.val) for the fixed effects of breed on factor score and performed ANOVA to obtain the breed F statistics. To obtain the likelihood ratio for each breed, we constructed models using maximum likelihood with and without the breed and performed an ANOVA.

GWASs using mixed linear models

We performed genome-wide mixed linear model–based associations in the Darwin's Ark genetic cohort using the "leave-one-chromosome-out" approach (MLMA-LOCO) implemented in GCTA (*56*) with categorical covariates for sex and data type (genotyping or low-pass sequencing) and quantitative covariates for height and age for nonmorphological traits. Because LD is nearly as short in diverse dogs as in humans, we used the thresholds for genome-wide significance ($p = 5 \times 10^{-8}$) and suggestive associations ($p = 1 \times 10^{-6}$) that are conventionally used in human GWASs (*1*, *76*).

We defined regions of association by clumping SNPs in LD ($r^2 > 0.2$ and $r^2 > 0.5$) and near (<250 kb) associated index SNPs using PLINK (data S16). When comparing region sizes to the earlier osteosarcoma study (*83*), we used the same clumping thresholds. To assess how much phenotypic variance was explained by associated regions, we derived genetic relationship matrices for regions of suggestive association ($p = 1 \times 10^{-6}$) with each trait and the set of all other SNPs and estimated the partitioned heritability as the proportion of total heritability unattributed by discovered associations.

We built a predictive model for height as responses to survey Q121 for 1730 dogs older than 18 months and assessed its power through 10-fold cross validation (9/10 training, 1/10 testing). At each round, we performed GWASs on the training set, selected SNPs for prediction at given p value cutoffs, built a random forest regression model, and assessed accuracy using the testing set. The reported accuracy and mean squared error are averaged across 10 rounds.

We tested for enrichment of association summary statistics in three types of gene sets (*22*) by applying MAGMA (version 1.09) (*91*), a method that accounts for region size, variant count, and LD (data S17).

Genetic differentiation of breeds

We calculated genome-wide normalized PBS scores using the Hudson estimator of fixation (F_{ST}) for each breed ($N_P \geq 12$ dogs, maximum 88) relative to dogs overall ($N_D = 3890 - N_P$) and wolves ($N_W = 48$) in sliding windows of 100 kb by 10 kb (data S18) over ~27.6 million SNPs from publicly available and Darwin's Ark genetic data (*22*). After dividing locus tests into physical trait, behavioral question, and behavioral factor associations, we performed a one-tailed Student's *t* test to test whether the observed maximum PBS within associated loci exceeded what we expect by random chance (data S20). To test whether allele frequencies at SNPs associated with behavioral or physical traits tended to differ more in breeds, we calculated the max F_{ST} observed between one of the top 10 breeds and all other dogs and compared this to 29,903 randomly sampled SNPs using a one-sided *t* test.

REFERENCES AND NOTES

1. K. Lindblad-Toh *et al*., Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* **438**, 803–819 (2005). doi: 10.1038/nature04338; pmid: 16341006
2. A. Bergström *et al*., Origins and genetic legacy of prehistoric dogs. *Science* **370**, 557–564 (2020). doi: 10.1126/science.aba9572; pmid: 33122379
3. K. Lord, M. Feinstein, B. Smith, R. Coppinger, Variation in reproductive traits of members of the genus *Canis* with special attention to the domestic dog (*Canis familiaris*). *Behav. Processes* **92**, 131–142 (2013). doi: 10.1016/j.beproc.2012.10.009; pmid: 23124015
4. C. Hansen Wheat, W. van der Bijl, H. Temrin, Dogs, but not wolves, lose their sensitivity toward novelty with age. *Front. Psychol.* **10**, 2001 (2019). doi: 10.3389/fpsyg.2019.02001; pmid: 31555182
5. L. Moretti, M. Hentrup, K. Kotrschal, F. Range, The influence of relationships on neophobia and exploration in wolves and dogs. *Anim. Behav.* **107**, 159–173 (2015). doi: 10.1016/j.anbehav.2015.06.008; pmid: 26405301
6. M. Arendt, K. M. Cairns, J. W. O. Ballard, P. Savolainen, E. Axelsson, Diet adaptation in dog reflects spread of prehistoric agriculture. *Heredity* **117**, 301–306 (2016). doi: 10.1038/hdy.2016.48; pmid: 27406651
7. E. Axelsson *et al*., The genomic signature of dog domestication reveals adaptation to a starch-rich diet. *Nature* **495**, 360–364 (2013). doi: 10.1038/nature11837; pmid: 23354050
8. J. de G. Mazzorin, A. Tagliacozzo, "Morphological and osteological changes in the dog from the Neolithic to the Roman period in Italy" in *Dogs Through Time: An Archaeological Perspective* (Archaeopress, 2000), pp. 141–161.
9. J. D. G. Mazzorin, A. Tagliacozzo, Dog remains in Italy from the Neolithic to the Roman period. *Anthropozoologica* **25**, 429 (1997).
10. M. Worboys, J.-M. Strange, N. Pemberton, *The Invention of the Modern Dog: Breed and Blood in Victorian Britain* (JHU Press, 2018).
11. H. Ritvo, Pride and pedigree: The evolution of the Victorian dog fancy. *Vic. Stud.* **29**, 227–253 (1986).
12. K. Lord, R. A. Schneider, R. Coppinger, in *The Domestic Dog: Its Evolution, Behavior and Interactions with People*, J. Serpell, Ed. (Cambridge Univ. Press, 2016), chap. 4, pp. 42–66.
13. C. D. Arons, W. J. Shoemaker, The distribution of catecholamines and β-endorphin in the brains of three behaviorally distinct breeds of dogs and their F1 hybrids. *Brain Res.* **594**, 31–39 (1992). doi: 10.1016/0006-8993(92)91026-B; pmid: 1467939
14. E. K. Karlsson *et al*., Efficient mapping of mendelian traits in dogs through genome-wide association. *Nat. Genet.* **39**, 1321–1328 (2007). doi: 10.1038/ng.2007.10; pmid: 17906626
15. E. K. Karlsson, K. Lindblad-Toh, Leader of the pack: Gene mapping in dogs and other model organisms. *Nat. Rev. Genet.* **9**, 713–725 (2008). doi: 10.1038/nrg2382; pmid: 18714291
16. N. B. Sutter, E. A. Ostrander, Dog star rising: The canine genetic system. *Nat. Rev. Genet.* **5**, 900–910 (2004). doi: 10.1038/nrg1492; pmid: 15573122
17. M. E. Gompper, Ed., *Free-Ranging Dogs and Wildlife Conservation* (Oxford Univ. Press, 2013).
18. American Veterinary Medical Association (AVMA), *United States Pet Ownership and Demographics Sourcebook* (AVMA, Veterinary Economics Division, 2018).
19. D. C. Coile, *Encyclopedia of Dog Breeds* (Barrons Educational Series, 1998).
20. American Kennel Club, *The New Complete Dog Book: Official Breed Standards and Profiles for Over 200 Breeds* (Fox Chapel Publishing, 2017).
21. J. P. Scott, J. L. Fuller, *Genetics and the Social Behavior of the Dog* (Univ. Chicago Press, 1965).
22. See the supplementary materials.
23. L. R. Mehrkam, C. D. L. Wynne, Behavioral differences among breeds of domestic dogs (*Canis lupus familiaris*): Current status of the science. *Appl. Anim. Behav. Sci.* **155**, 12–27 (2014). doi: 10.1016/j.applanim.2014.03.005
24. K. Svartberg, Breed-typical behaviour in dogs—Historical remnants or recent constructs? *Appl. Anim. Behav. Sci.* **96**, 293–313 (2006). doi: 10.1016/j.applanim.2005.06.014
25. H. J. Noh *et al*., Integrating evolutionary and regulatory information with a multispecies approach implicates genes and pathways in obsessive-compulsive disorder. *Nat. Commun.* **8**, 774 (2017). doi: 10.1038/s41467-017-00831-x; pmid: 29042551
26. N. H. Dodman *et al*., A canine chromosome 7 locus confers compulsive disorder susceptibility. *Mol. Psychiatry* **15**, 8–10 (2010). doi: 10.1038/mp.2009.111; pmid: 20029408
27. K. L. Overall, Natural animal models of human psychiatric conditions: Assessment of mechanism and validity. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **24**, 727–776 (2000). doi: 10.1016/S0278-5846(00)00104-4; pmid: 11191711
28. R. Sarviaho *et al*., A novel genomic region on chromosome 11 associated with fearfulness in dogs. *Transl. Psychiatry* **10**, 169 (2020). doi: 10.1038/s41398-020-0849-z; pmid: 32467585
29. R. Sarviaho *et al*., Two novel genomic regions associated with fearfulness in dogs overlap human neuropsychiatric loci. *Transl. Psychiatry* **9**, 18 (2019). doi: 10.1038/s41398-018-0361-x; pmid: 30655508
30. J. Ilska *et al*., Genetic characterization of dog personality traits. *Genetics* **206**, 1101–1111 (2017). doi: 10.1534/genetics.116.192674; pmid: 28396505
31. R. Tang *et al*., Candidate genes and functional noncoding variants identified in a canine model of obsessive-compulsive disorder. *Genome Biol.* **15**, R25 (2014). doi: 10.1186/gb-2014-15-3-r25; pmid: 24995881
32. C. DeBoever *et al*., Assessing digital phenotyping to enhance genetic studies of human diseases. *Am. J. Hum. Genet.* **106**, 611–622 (2020). doi: 10.1016/j.ajhg.2020.03.007; pmid: 32275883
33. C. L. Hyde *et al*., Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. *Nat. Genet.* **48**, 1031–1036 (2016). doi: 10.1038/ng.3623; pmid: 27479909
34. H. F. Wright, D. S. Mills, P. Pollux, Development and validation of a psychometric tool for assessing impulsivity in the domestic dog, *Canis familiaris*. *Int. J. Comp. Psychol.* (2011). doi: 10.1016/j.physbeh.2011.09.019; pmid: 21986321
35. H. E. Salvin, P. D. McGreevy, P. S. Sachdev, M. J. Valenzuela, The canine cognitive dysfunction rating scale (CCDR): A data-driven and ecologically relevant assessment tool. *Vet. J.* **188**, 331–336 (2011). doi: 10.1016/j.tvjl.2010.05.014; pmid: 20542455
36. R. P. Lavan, Development and validation of a survey for quality of life assessment by owners of healthy dogs. *Vet. J.* **197**, 578–582 (2013). doi: 10.1016/j.tvjl.2013.03.021; pmid: 23639368
37. A. C. Jones, *Development and Validation of a Dog Personality Questionnaire* (Univ. Texas, 2008).
38. N. B. Sutter, D. S. Mosher, M. M. Gray, E. A. Ostrander, Morphometrics within dog breeds are highly reproducible and dispute Rensch's rule. *Mamm. Genome* **19**, 713–723 (2008). doi: 10.1007/s00335-008-9153-6; pmid: 19020935
39. J. Yordy *et al*., Body size, inbreeding, and lifespan in domestic dogs. *Conserv. Genet.* **21**, 137–148 (2020). doi: 10.1007/s10592-019-01240-x; pmid: 32607099
40. J. Plassais *et al*., Whole genome sequencing of canids reveals genomic regions under selection and variants influencing morphology. *Nat. Commun.* **10**, 1489 (2019). doi: 10.1038/s41467-019-09373-w; pmid: 30940804
41. H. G. Parker *et al*., Genomic analyses reveal the influence of geographic origin, migration, and hybridization on modern dog breed development. *Cell Rep.* **19**, 697–708 (2017). doi: 10.1016/j.celrep.2017.03.079; pmid: 28445722
42. A. R. Martin *et al*., Low-coverage sequencing cost-effectively detects known and novel variation in underrepresented populations. *Am. J. Hum. Genet.* **108**, 656–668 (2021). doi: 10.1016/j.ajhg.2021.03.012; pmid: 33770507
43. S. Rubinacci, D. M. Ribeiro, R. J. Hofmeister, O. Delaneau, Efficient phasing and imputation of low-coverage sequencing data using large reference panels. *Nat. Genet.* **53**, 120–126 (2021). doi: 10.1038/s41588-020-00756-0; pmid: 33414550
44. R. M. Buckley *et al*., Best practices for analyzing imputed genotypes from low-pass sequencing in dogs. *Mamm. Genome* **33**, 213–229 (2021). doi: 10.1007/s00335-021-09914-z; pmid: 34498136
45. J. H. Li, C. A. Mazur, T. Berisa, J. K. Pickrell, Low-pass sequencing increases the power of GWAS and decreases measurement error of polygenic risk scores compared to genotyping arrays. *Genome Res.* **31**, 529–537 (2021). doi: 10.1101/gr.266486.120; pmid: 33536225
46. K. Wasik *et al*., Comparing low-pass sequencing and genotyping for trait mapping in pharmacogenetics. *BMC Genomics* **22**, 197 (2021). doi: 10.1186/s12864-021-07508-2; pmid: 33743587
47. D. H. Alexander, K. Lange, Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics* **12**, 246 (2011). doi: 10.1186/1471-2105-12-246; pmid: 21682921
48. J. Yang *et al*., Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **42**, 565–569 (2010). doi: 10.1038/ng.608; pmid: 20562875
49. G. Athanasiadis *et al*., Estimating narrow-sense heritability using family data from admixed populations. *Heredity* **124**, 751–762 (2020). doi: 10.1038/s41437-020-0311-2; pmid: 32273574
50. S. Luo *et al*., Genome-wide association study of serum metabolites in the African American Study of Kidney Disease and Hypertension. *Kidney Int.* **100**, 430–439 (2021). doi: 10.1016/j.kint.2021.03.026; pmid: 33838163
51. R. Bakeman, Recommended effect size statistics for repeated measures designs. *Behav. Res. Methods* **37**, 379–384 (2005). doi: 10.3758/BF03192707; pmid: 16405133
52. American Kennel Club, Dog breeds; https://www.akc.org/dog-breeds/.
53. American Kennel Club, *The Complete Dog Book* (Ballantine Books, 2006).
54. D. L. Wells, D. J. Morrison, P. G. Hepper, The effect of priming on perceptions of dog breed traits. *Anthrozoos* **25**, 369–377 (2012). doi: 10.2752/175303712X13403555186370
55. L. M. Gunter, R. T. Barber, C. D. L. Wynne, What's in a name? Effect of breed perceptions & labeling on attractiveness, adoptions & length of stay for pit-bull-type dogs. *PLOS ONE* **11**, e0146857 (2016). doi: 10.1371/journal.pone.0146857; pmid: 27008213
56. J. Yang, S. H. Lee, M. E. Goddard, P. M. Visscher, GCTA: A tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011). doi: 10.1016/j.ajhg.2010.11.011; pmid: 21167468
57. M. Pirinen, P. Donnelly, C. C. A. Spencer, Efficient computation with a linear mixed model on large-scale data sets with applications to genetic studies. *Ann. Appl. Stat.* **7**, 369–390 (2013). doi: 10.1214/12-AOAS586
58. E. Cadieu *et al*., Coat variation in the domestic dog is governed by variants in three genes. *Science* **326**, 150–153 (2009). doi: 10.1126/science.1177808; pmid: 19713490
59. L. Brancalion *et al*., Roan, ticked and clear coat patterns in the canine are associated with three haplotypes near *usherin* on CFA38. *Anim. Genet.* **52**, 198–207 (2021). doi: 10.1111/age.13040; pmid: 33539602
60. A. J. Slavney *et al*., Five genetic variants explain over 70% of hair coat pheomelanin intensity variation in purebred and mixed breed domestic dogs. *PLOS ONE* **16**, e0250579 (2021). doi: 10.1371/journal.pone.0250579; pmid: 34043658
61. S. I. Candille *et al*., A β-defensin mutation causes black coat color in domestic dogs. *Science* **318**, 1418–1423 (2007). doi: 10.1126/science.1147880; pmid: 17947548
62. M. Oguro-Okano, M. Honda, K. Yamazaki, K. Okano, Mutations in the *melanocortin 1 receptor*, *β-defensin103* and *agouti signaling protein* genes, and their association with coat color phenotypes in Akita-inu dogs. *J. Vet. Med. Sci.* **73**, 853–858 (2011). doi: 10.1292/jvms.10-0439; pmid: 21321476

63. J. A. Kerns *et al*., Linkage and segregation analysis of black and brindle coat color in domestic dogs. *Genetics* **176**, 1679–1689 (2007). doi: 10.1534/genetics.107.074237; pmid: 17483404
64. K. Chase *et al*., Genetic basis for systems of skeletal quantitative traits: Principal component analysis of the canid skeleton. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 9930–9935 (2002). doi: 10.1073/pnas.152333099; pmid: 12114542
65. K. Chase, D. R. Carrier, F. R. Adler, E. A. Ostrander, K. G. Lark, Interaction between the X chromosome and an autosome regulates size sexual dimorphism in Portuguese water dogs. *Genome Res.* **15**, 1820–1824 (2005). doi: 10.1101/gr.3712705; pmid: 16339380
66. N. B. Sutter *et al*., A single *IGF1* allele is a major determinant of small size in dogs. *Science* **316**, 112–115 (2007). doi: 10.1126/science.1137045; pmid: 17412960
67. H. G. Parker *et al*., An expressed *fgf4* retrogene is associated with breed-defining chondrodysplasia in domestic dogs. *Science* **325**, 995–998 (2009). doi: 10.1126/science.1173275; pmid: 19608863
68. B. C. Hoopes, M. Rimbault, D. Liebers, E. A. Ostrander, N. B. Sutter, The insulin-like growth factor 1 receptor (IGF1R) contributes to reduced size in dogs. *Mamm. Genome* **23**, 780–790 (2012). doi: 10.1007/s00335-012-9417-z; pmid: 22903739
69. M. Rimbault *et al*., Derived variants at six genes explain nearly half of size reduction in dog breeds. *Genome Res.* **23**, 1985–1995 (2013). doi: 10.1101/gr.157339.113; pmid: 24026177
70. E. A. Brown *et al*., *FGF4* retrogene on CFA12 is responsible for chondrodystrophy and intervertebral disc disease in dogs. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 11476–11481 (2017). doi: 10.1073/pnas.1709082114; pmid: 29073074
71. M. T. Webster *et al*., Linked genetic variants on chromosome 10 control ear morphology and body mass among dog breeds. *BMC Genomics* **16**, 474 (2015). doi: 10.1186/s12864-015-1702-2; pmid: 26100605
72. T. Okada *et al*., Anderson's disease/chylomicron retention disease in a Japanese patient with uniparental disomy 7 and a normal *SAR1B* gene protein coding sequence. *Orphanet J. Rare Dis.* **6**, 78 (2011). doi: 10.1186/1750-1172-6-78; pmid: 22104167
73. E. Levy *et al*., Sar1b transgenic male mice are more susceptible to high-fat diet-induced obesity, insulin insensitivity and intestinal chylomicron overproduction. *J. Nutr. Biochem.* **25**, 540–548 (2014). doi: 10.1016/j.jnutbio.2014.01.004; pmid: 24657056
74. N. F. Ajeawung *et al*., Mutations in *ANAPC1*, encoding a scaffold subunit of the anaphase-promoting complex, cause Rothmund-Thomson syndrome type 1. *Am. J. Hum. Genet.* **105**, 625–630 (2019). doi: 10.1016/j.ajhg.2019.06.011; pmid: 31303264
75. T. L. Gumienny *et al*., *Caenorhabditis elegans* SMA-10/LRIG is a conserved transmembrane protein that enhances bone morphogenetic protein signaling. *PLOS Genet.* **6**, e1000963 (2010). doi: 10.1371/journal.pgen.1000963; pmid: 20502686
76. Z. Chen, M. Boehnke, X. Wen, B. Mukherjee, Revisiting the genome-wide significance threshold for common variant GWAS. *G3* **11**, jkaa056 (2021). doi: 10.1093/g3journal/jkaa056; pmid: 33585870
77. J. J. Lee *et al*., Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat. Genet.* **50**, 1112–1121 (2018). doi: 10.1038/s41588-018-0147-3; pmid: 30038396
78. J. E. Savage *et al*., Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nat. Genet.* **50**, 912–919 (2018). doi: 10.1038/s41588-018-0152-6; pmid: 29942086
79. M. Lam *et al*., Pleiotropic meta-analysis of cognition, education, and schizophrenia differentiates roles of early neurodevelopmental and adult synaptic pathways. *Am. J. Hum. Genet.* **105**, 334–350 (2019). doi: 10.1016/j.ajhg.2019.06.012; pmid: 31374203
80. R. S. Smith *et al*., Sodium channel SCN3A ($Na_V1.3$) regulation of human cerebral cortical folding and oral motor development. *Neuron* **99**, 905–913.e7 (2018). doi: 10.1016/j.neuron.2018.07.052; pmid: 30146301
81. Y. F. Widmer, A. Bilican, R. Bruggmann, S. G. Sprecher, Regulators of long-term memory revealed by mushroom body-specific gene expression profiling in *Drosophila melanogaster*. *Genetics* **209**, 1167–1181 (2018). doi: 10.1534/genetics.118.301106; pmid: 29925565
82. M. Pelé, L. Tiret, J.-L. Kessler, S. Blot, J.-J. Panthier, SINE exonic insertion in the *PTPLA* gene leads to multiple splicing defects and segregates with the autosomal recessive centronuclear myopathy in dogs. *Hum. Mol. Genet.* **14**, 1417–1427 (2005). doi: 10.1093/hmg/ddi151; pmid: 15829503
83. E. K. Karlsson *et al*., Genome-wide analyses implicate 33 loci in heritable dog osteosarcoma, including regulatory variants near *CDKN2A/B*. *Genome Biol.* **14**, R132 (2013). doi: 10.1186/gb-2013-14-12-r132; pmid: 24330828
84. A. R. Boyko *et al*., A simple genetic architecture underlies morphological variation in dogs. *PLOS Biol.* **8**, e1000451 (2010). doi: 10.1371/journal.pbio.1000451; pmid: 20711490
85. J. Lonsdale *et al*., The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45**, 580–585 (2013). doi: 10.1038/ng.2653; pmid: 23715323
86. E. L. MacLean, N. Snyder-Mackler, B. M. vonHoldt, J. A. Serpell, Highly heritable and functionally relevant breed differences in dog behaviour. *Proc. Biol. Sci.* **286**, 20190716 (2019). doi: 10.1098/rspb.2019.0716; pmid: 31575369
87. A. P. Privitera *et al*., OCDB: A database collecting genes, miRNAs and drugs for obsessive-compulsive disorder. *Database* **2015**, bav069 (2015). doi: 10.1093/database/bav069; pmid: 26228432
88. B. S. Abrahams *et al*., SFARI Gene 2.0: A community-driven knowledgebase for the autism spectrum disorders (ASDs). *Mol. Autism* **4**, 36 (2013). doi: 10.1186/2040-2392-4-36; pmid: 24090431
89. Schizophrenia Working Group of the Psychiatric Genomics Consortium, Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427 (2014). doi: 10.1038/nature13595; pmid: 25056061
90. A. F. Pardiñas *et al*., Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nat. Genet.* **50**, 381–389 (2018). doi: 10.1038/s41588-018-0059-2; pmid: 29483656
91. C. A. de Leeuw, J. M. Mooij, T. Heskes, D. Posthuma, MAGMA: Generalized gene-set analysis of GWAS data. *PLOS Comput. Biol.* **11**, e1004219 (2015). doi: 10.1371/journal.pcbi.1004219; pmid: 25885710
92. X. Yi *et al*., Sequencing of 50 human exomes reveals adaptation to high altitude. *Science* **329**, 75–78 (2010). doi: 10.1126/science.1190371; pmid: 20595611
93. B. vonHoldt, Z. Fan, D. Ortega-Del Vecchyo, R. K. Wayne, *EPAS1* variants in high altitude Tibetan wolves were selectively introgressed into highland dogs. *PeerJ* **5**, e3522 (2017). doi: 10.7717/peerj.3522; pmid: 28717592
94. M. S. Sinding *et al*., Arctic-adapted dogs emerged at the Pleistocene-Holocene transition. *Science* **368**, 1495–1499 (2020). doi: 10.1126/science.aaz8599; pmid: 32587022
95. P. E. Deane-Coe, E. T. Chu, A. Slavney, A. R. Boyko, A. J. Sams, Direct-to-consumer DNA testing of 6,000 dogs reveals 98.6-kb duplication associated with blue eyes and heterochromia in Siberian Huskies. *PLOS Genet.* **14**, e1007648 (2018). doi: 10.1371/journal.pgen.1007648; pmid: 30286082
96. J. S. Sanjak, J. Sidorenko, M. R. Robinson, K. R. Thornton, P. M. Visscher, Evidence of directional and stabilizing selection in contemporary humans. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 151–156 (2018). doi: 10.1073/pnas.1707227114; pmid: 29255044
97. P. C. Sabeti *et al*., Positive natural selection in the human lineage. *Science* **312**, 1614–1620 (2006). doi: 10.1126/science.1124309; pmid: 16778047
98. E. G. Atkinson *et al*., Tractor uses local ancestry to enable the inclusion of admixed individuals in GWAS and to boost power. *Nat. Genet.* **53**, 195–204 (2021). doi: 10.1038/s41588-020-00766-y; pmid: 33462486
99. N. Barton, J. Hermisson, M. Nordborg, Population genetics: Why structure matters. *eLife* **8**, e45380 (2019). doi: 10.7554/eLife.45380; pmid: 30895925
100. G. L. Wojcik *et al*., Genetic analyses of diverse populations improves discovery for complex traits. *Nature* **570**, 514–518 (2019). doi: 10.1038/s41586-019-1310-4; pmid: 31217584
101. J. E. Moore *et al*., Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature* **583**, 699–710 (2020). doi: 10.1038/s41586-020-2493-4; pmid: 32728249
102. K. Megquier *et al*., BarkBase: Epigenomic annotation of canine genomes. *Genes* **10**, 433 (2019). doi: 10.3390/genes10060433; pmid: 31181663
103. Zoonomia Consortium, A comparative genomics multitool for scientific discovery and conservation. *Nature* **587**, 240–245 (2020). doi: 10.1038/s41586-020-2876-6; pmid: 33177664
104. K. Morrill *et al*., Associated data files and scripts for "Ancestry-inclusive dog genomics challenges popular breed stereotypes". *Dryad* (2022); doi: 10.5061/dryad.g4f4qrfr0
105. K. Morrill *et al*., Associated data files and scripts for "Ancestry-inclusive dog genomics challenges popular breed stereotypes". *Zenodo* (2022); doi: 10.5281/zenodo.5808330
106. K. Morrill, Darwin's Ark - Data release 2022. Terra (2022); https://app.terra.bio/#workspaces/darwins-ark/Darwins%20Ark%20-%20Data%20Release%202022.
107. K. Morrill *et al*., Additional figures for "Ancestry-inclusive dog genomics challenges popular breed stereotypes". *FigShare* (2022); doi: 10.6084/m9.figshare.16608793

ACKNOWLEDGMENTS

We thank all of our participants and their dogs for making this work possible, as well as the International Association of Animal Behavior Consultants community for their feedback and help with outreach; E. A. Ostrander for helpful advice and the Ostrander lab for providing the high-coverage whole-genome data (available in data S4); B. Klein and the National Entlebucher Mountain Dog Association for providing whole-genome data; R. Peters and the 360design.com team; J. Gallagher, H. Herzog, A. Karlsson, K. Lindblad-Toh, L. Moses, B. Neale, C. Painter, D. Promislow, B. Rosener, and D. Weaver for useful discussions; H. Beberman for support; B. Burton, R. Crisler, G. Fisher, S. Fraser, L. Haug, J. Hourihan, M. Mullins, C. Pachel, L. Strassberg, and M. Workman for input on survey development; C. Abrams, R. Bacon, K. Bigger, M. Bishop, J. Quintal, D. Church, J. Conner, A. Dauphin, A. Derr, T. and K. Flotte, T. Fortier, S. Garamszegi, B. R. Granger, S. Humphreys, C. Mitchell, M. Movassagh, A. Jorgensen, M. Lane, A. and M. Lek, B. Lengsfeld, J. Luban, A. Macbeth, L. McGuire, S. and J. Moeling, D. Morrissey, E. Neenan, J. O'Donnell, T. Ollier, L. Paquin, G. Peloso, A. Pensarosa, A. Phelps, S. Richardson, J. Rivera, S. Rulnick, S. Schnaffner, R. Skloot, E. Stackpole, W. Theurkauf, S. Thier, and E. Winchester for their time and thoughtful replies to our many questions about their dogs as well as permission to share photographs of their dogs. **Funding:** This project was funded by the National Institutes of Health, including NIMH R21 MH109938, NCI R01 CA255319, NCI R37 CA218570, NCI F32 CA247088, NHGRI R01 HG008742, NHGRI U24 HG009446, OD R24 OD018250, and NIA U19 AG057377. It was also supported by NSF EF-2022007, Broad Institute BroadIgnite and Next10 awards, the Darwin's Ark Foundation, the Food Allergy Science Initiative, the Manton Foundation, and the Working Dog Project. **Author contributions:** E.K.K. conceived of the project; K.Mo., J.M., M.A., H.J.N., C.L., K.A.L., and E.K.K. designed the project; K.Mo., J.H., X.L., J.M., B.L., L.G., M.G., M.A., E.C., J.A., J.J., M.K., C.L., K.Me., R.S., J.T.-M., D.P.G., K.A.L., and E.K.K. performed data acquisition; K.Mo., J.H., X.L., J.M., B.L., L.G., M.G., E.C., N.S.-M., H.J.N., K.Me., R.S., J.T.-M., Z.W., D.P.G., K.A.L., and E.K.K. performed analyses; K.Mo., J.H., J.M., L.G., Y.D., and A.C. developed software; K.Mo., J.H., X.L., B.L., D.P.G., K.A.L., and E.K.K. wrote the manuscript; and J.H., M.A., N.S.-M., H.J.N., C.L., K.Me., and M.E.W. edited the manuscript. **Competing interests:** L.G. is a co-founder of, equity owner in, and chief technical officer at Fauna Bio Inc. H.J.N is an employee of AbbVie Inc. **Data and materials availability:** All sequencing data have been deposited into the National Center for Biotechnology Information (NCBI) Sequence Read Archive, including high-coverage sequencing reads for the Mendel's Mutts dataset in BioProject PRJNA683923 and low-coverage sequencing reads deposited to BioProject PRJNA675863. The Broad-UMass Canid Variants in variant call format (VCF) is available publicly via FTP at https://data.broadinstitute.org/DogData/. All survey and genetic data from dogs of the Darwin's Ark and Mendel's Mutts cohorts, data from the MuttMix survey, and scripts used in analyses are archived in Dryad (*104*) and Zenodo (*105*). The genome-wide association summary statistics and plots are shared on Terra (*106*) and Figshare (*107*), respectively.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abk0639
Materials and Methods
Supplementary Text
Figs. S1 to S34
Tables S1 to S15
References (*108*–*169*)
MDAR Reproducibility Checklist
Data S1 to S20

Submitted 6 October 2021; accepted 22 March 2022
10.1126/science.abk0639

Science
JOURNALS AAAS

Publish your research in the *Science* family of journals

The *Science* family of journals (*Science, Science Advances, Science Immunology, Science Robotics, Science Signaling*, and *Science Translational Medicine)* are among the most highly-regarded journals in the world for quality and selectivity. Our peer-reviewed journals are committed to publishing cutting-edge research, incisive scientific commentary, and insights on what's important to the scientific world at the highest standards.

Submit your research today!

Learn more at **Science.org/journals**

RESEARCH ARTICLES

CANCER

Cancer cells use self-inflicted DNA breaks to evade growth limits imposed by genotoxic stress

Brian D. Larsen[1], Jan Benada[1], Philip Yuk Kwong Yung[2]†, Ryan A. V. Bell[3]†, George Pappas[4]†, Vaclav Urban[5], Johanna K. Ahlskog[1], Tia T. Kuo[1], Pavel Janscak[5,6], Lynn A. Megeney[3], Simon J. Elsässer[2], Jiri Bartek[2,4,5], Claus S. Sørensen[1]*

Genotoxic therapy such as radiation serves as a frontline cancer treatment, yet acquired resistance that leads to tumor reoccurrence is frequent. We found that cancer cells maintain viability during irradiation by reversibly increasing genome-wide DNA breaks, thereby limiting premature mitotic progression. We identify caspase-activated DNase (CAD) as the nuclease inflicting these de novo DNA lesions at defined loci, which are in proximity to chromatin-modifying CCCTC-binding factor (CTCF) sites. CAD nuclease activity is governed through phosphorylation by DNA damage response kinases, independent of caspase activity. In turn, loss of CAD activity impairs cell fate decisions, rendering cancer cells vulnerable to radiation-induced DNA double-strand breaks. Our observations highlight a cancer-selective survival adaptation, whereby tumor cells deploy regulated DNA breaks to delimit the detrimental effects of therapy-evoked DNA damage.

Genotoxic cancer therapy inactivates and kills cancer cells by inflicting extensive DNA damage. Radiation therapy (RT) is the most broadly applied genotoxic challenge in standard-of-care oncological treatment. The deposition of energy as radiation passes through the genetic material triggers extensive DNA lesions often in the form of double-strand breaks (DSBs), single-stranded breaks (SSBs), and interstrand cross-links (*1*). The extent of this damage can present an insurmountable barrier to cellular fitness, triggering cell death or cell cycle withdrawal. Facilitating the DNA damage response (DDR) and lesion repair while avoiding cell death and cell cycle blockage is critical for cell survival after RT. Clinically, resistance to RT remains a considerable obstacle to effective tumor control, as the cancer cells deploy an arsenal of mechanisms, still incompletely understood, to mitigate the effects of RT (*1*).

Irradiated normal cells halt progression in the G_1 phase of the cell cycle by activation of p53 and pRB, which are key factors regulating cell cycle checkpoints. However, these factors are commonly inactivated in solid cancers, leading to G_1 checkpoint deficiency. Combined with oncogene-driven premature S-phase entry, this scenario evokes replication stress and enhanced chromosomal damage that requires efficient repair should the cell yield viable progenies when it divides (*2*). Hence, cancer cells particularly rely on the G_2 cell cycle checkpoint, preventing entry into mitosis with unrepaired DNA breaks (*3*). In addition, studies following the dynamics of RT-induced DNA lesions have indicated the presence of a temporally distinct and unexplained secondary wave of DNA breaks (*4*). To identify potential nuclease regulators of the G_2 cell cycle checkpoint after radiation, we screened a library targeting the known human nucleases in cancer cells (fig. S1) (*5*). The primary candidate emerging from this screen was RBBP8 (CtIP), an established DDR and G_2 checkpoint factor (*6*). Unexpectedly, a second robust target from this screen was DFFB (also known as caspase-activated DNase or CAD), a factor previously unassociated with DDR or cell cycle checkpoint control. CAD is the nuclease implicated in DNA fragmentation during apoptotic cell death as the effector of caspase signaling cascades (*7*, *8*). Caspase-mediated cleavage of CAD's chaperone and inhibitor, ICAD (inhibitor of CAD), facilitates the dimerization of CAD, giving rise to the hallmark DNA fragmentation seen in apoptosis (*8*).

CAD-inflicted DNA breaks and the ensuing DDR signals have also been implicated as inductive cues for a number of nonapoptotic cell fate states and transitions (*9*–*15*). Intriguingly, RT-induced DNA lesions encompass a temporally distinct and mechanistically unexplained secondary wave of DNA breaks (*4*). Together, these observations led us to hypothesize that CAD might be responsible for inflicting these delayed post-irradiation DNA breaks to exert checkpoint control, potentially providing a mechanism of radioresistance.

CAD promotes a wave of endogenous DNA breaks after exogenous DNA damage

To address the nature of the endogenous DNA breaks that have been reported to appear after exposure to ionizing radiation (IR), we irradiated human wild-type and CAD null (KO) colorectal cancer–derived HCT116 cells and measured the extent of DNA damage by alkaline single-cell gel electrophoresis (Comet assay) (Fig. 1A and fig. S2A). We did not observe any initial differences in DNA lesion accumulation between wild-type and CAD KO cells after exposure to IR. Further, the progressive reduction in DNA damage burden through active DNA repair was comparable between wild-type and CAD KO cells. However, consistent with recent observations (*4*), we detected a secondary accumulation of DNA lesions in wild-type cells that was prominent 24 hours after IR (Fig. 1A). The extent of DNA breakage was dependent on IR dosage (Fig. 1B). DNA break accumulation at 24 hours was largely dependent on the nuclease activity of CAD (Fig. 1C). These CAD-dependent breaks were observed in a panel of human malignant cell lines, but not in cells of nonmalignant origin (Fig. 1D and fig. S2B). To support this finding, we used in situ nick translation (ISNT) to quantify DNA breaks (*15*). This approach revealed a similar CAD-dependent elevation in DNA breaks at 24 hours after IR (Fig. 1E). CAD is normally present in a protein complex with ICAD, which also acts as a chaperone required to properly fold CAD (*8*). Accordingly, we recapitulated our observations in cells lacking ICAD, where the expression of CAD is lost (fig. S2, C and D). Further, extensive DNA DSB formation could not account for the prevalent CAD-dependent DNA breaks seen by the alkaline approach, although a discrete population of DNA DSBs might exist (fig. S2E).

PARP-1 [poly(ADP-ribose)polymerase–1] serves as a sensor of DNA lesions that triggers DNA repair and was previously implicated in the repair of DNA breaks after IR (*4*, *16*). Consistent with the appearance of DNA breaks at 24 hours after IR, we observed active ongoing DNA repair signaling, as evident from elevated poly(ADP-ribose) quantities (fig. S3A). Inhibition of PARP with 4-ANI, a strong enzymatic-activity inhibitor with only weak PARP-trapping activity, led to an elevation in DNA breaks even when the inhibitor was added for the last 2 hours of the 24-hour period after IR (fig. S3B) (*17*). The elevated PARP activity and increased DNA breaks upon PARP inhibition were both dependent on CAD (fig. S3, A and B).

The extent of single-stranded DNA (ssDNA) at 24 hours after IR was also monitored by

[1]Biotech Research and Innovation Centre, University of Copenhagen, 2200 N Copenhagen, Denmark. [2]Science for Life Laboratory, Division of Genome Biology, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, 17165 Stockholm, Sweden. [3]Sprott Centre for Stem Cell Research, Ottawa Hospital Research Institute and Departments of Medicine and Cellular and Molecular Medicine, University of Ottawa, Ottawa, Ontario K1H 8L6, Canada. [4]Danish Cancer Society Research Center, 2100 Copenhagen, Denmark. [5]Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, 143 00 Prague, Czech Republic. [6]Institute of Molecular Cancer Research, University of Zurich, 8057 Zurich, Switzerland.
*Corresponding author. Email: claus.storgaard@bric.ku.dk
†These authors contributed equally to this work.

native bromodeoxyuridine (BrdU) and replication protein A (RPA) foci formation. Both approaches demonstrated elevated numbers of ssDNA foci in the control irradiated cells but not in cells deficient for CAD/ICAD at this time point (fig. S3, C to H). In support of this observation, RPA foci formation at 24 hours could be restored by transient expression of wild-type CAD but not a nuclease-dead (ND) CAD variant (fig. S3I).

We postulated that elevation of DNA break quantities could signal a delayed chromatin response after IR. Consistent with this hypothesis, the maintenance of KAP1 phosphorylation, a chromatin marker of ongoing DDR, was dependent on CAD (fig. S4, A to C) (*18*). In addition to IR, we have observed a similar CAD-dependent signaling after doxorubicin-induced genotoxic damage, which also inflicts DNA DSBs (fig. S4, D to F). Collectively, these observations suggest that CAD nuclease induces self-inflicted DNA breaks in cancer cells. Canonically, caspase-3–mediated cleavage of ICAD releases CAD from inhibition, allowing for CAD dimerization. This positions each nuclease cleft in parallel across a DNA double-strand segment. Each nuclease cleft creates a single-stranded break in DNA that together produce a DSB (*8*). Notably, CAD can also inflict DNA nicks during early apoptosis and skeletal muscle differentiation (*19*, *20*). However, it has been reported that IR exposure of solid tumor cells does not elicit a robust caspase response (*21*). Here, we did not observe evidence of caspase-3 activation or proteolytic processing of ICAD after IR. Further, inhibition of pan-caspase activity did not affect the observed induction of DNA breaks, suggesting a noncanonical activation of CAD after IR (fig. S4, C, G, H, and I).

Chromatin recruitment of CAD and ICAD

We observed that both CAD and ICAD were recruited to the chromatin fraction of IR-treated cells (Fig. 2A and fig. S5A). ICAD interaction with CAD typically limits recruitment of the nuclease to DNA (*22*). However, the nuclease cleft of CAD is exposed in the CAD/ICAD heterodimer; thus, chromatin interactions could produce DNA nicks (*22*). To address whether chromatin recruitment of CAD/ICAD was sufficient to induce DNA breaks, we used the chromatin tethering model of U2OS 263 cells that harbors an integrated LacO array (*23*). Expressed mCherry-LacR-ICAD was correctly recruited to the LacO array and could recruit a green fluorescent protein (GFP)–tagged CAD to this site (fig. S5B). Analysis of the mCherry-LacR-ICAD tethered arrays in comparison to the empty mCherry-LacR construct revealed an induction of DNA breaks, as characterized by the creation of 3′-OH DNA ends that could be readily detected by terminal deoxynucleotidyl transferase end labeling

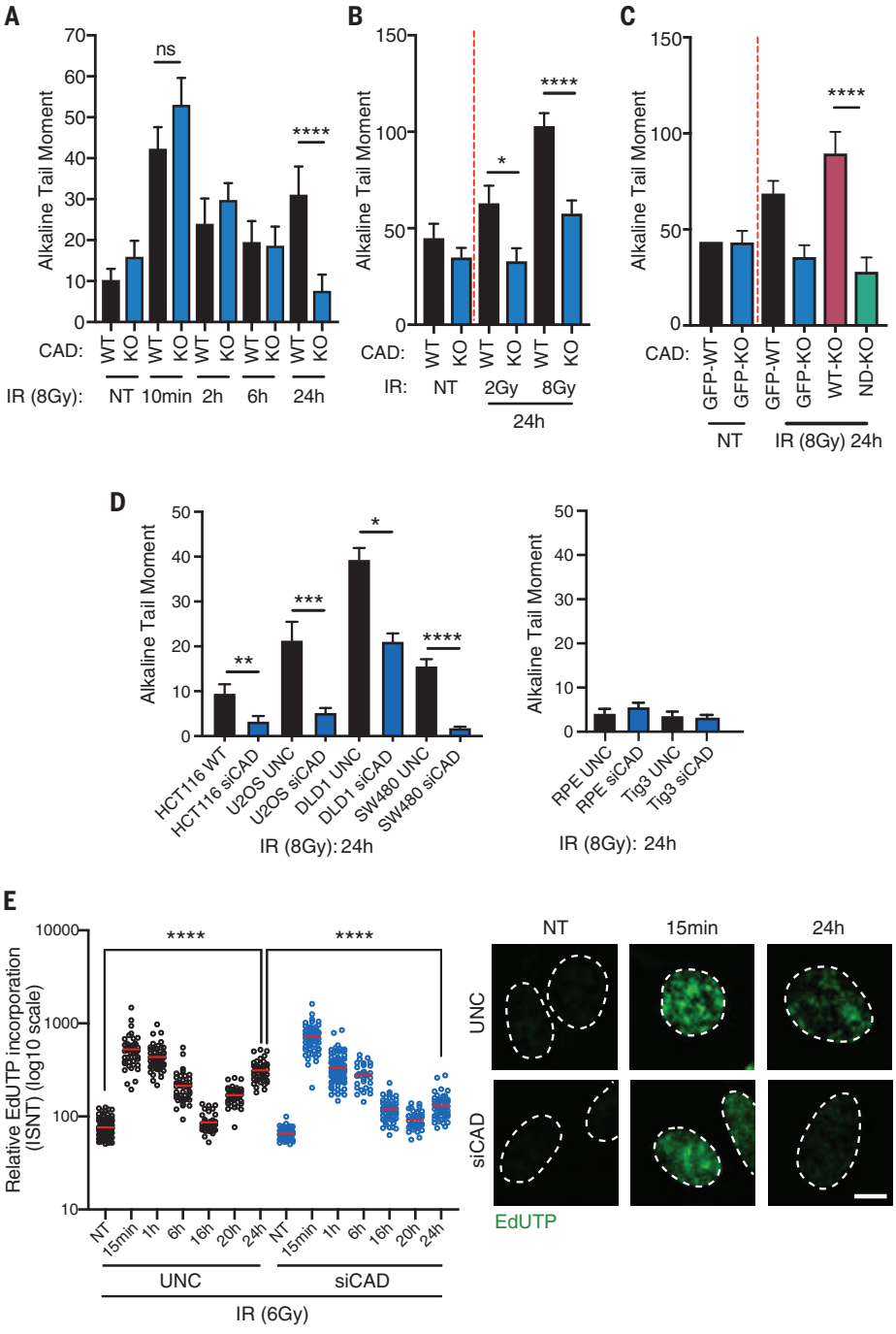


Fig. 1. CAD promotes self-inflicted DNA breaks after IR. (**A**) Alkaline Comet assay of HCT116 wild-type (WT) and CAD KO cells upon 8 Gy of IR; a representative dataset is presented. Data are means ± SEM; $N = 3$, $n > 100$. ****$P < 0.0001$ (Kruskal-Wallis multiple-comparisons test); ns, not significant. (**B**) Alkaline Comet assay of HCT116 WT and CAD KO cells upon 2 or 8 Gy of IR; a representative dataset is presented. Data are means ± SEM; $N = 3$, $n > 100$. *$P = 0.0147$, ****$P < 0.0001$ (Kruskal-Wallis multiple-comparisons test). (**C**) WT CAD, but not a ND CAD, restores DNA breaks as measured by the Alkaline Comet assay 24 hours after 8 Gy of IR; a representative dataset is presented. Data are means ± SEM; $N = 3$, $n > 100$. ****$P < 0.0001$ (Kruskal-Wallis multiple-comparisons test). (**D**) DNA break quantities 24 hours after 8 Gy of IR in cancer cell lines HCT116, U2OS, DLD-1, and SW480 and in noncancer cell lines RPE1 and Tig3 in control (UNC) or CAD-depleted cells (siCAD); a representative dataset is presented. Data are means ± SEM; $N = 3$, $n > 100$. *$P = 0.0337$, **$P = 0.074$, ***$P = 0.0001$, ****$P < 0.0001$ (Kruskal-Wallis multiple-comparisons test). (**E**) Relative ethynyl–deoxyuridine triphosphate incorporation into DNA breaks using ISNT in U2OS (control or CAD-depleted) cells upon 6 Gy of IR; a representative dataset is presented. Scale bar, 10 μm. Data are means ± SEM; $N = 3$, $n > 30$. ***$P < 0.001$ (unpaired Student's t test).

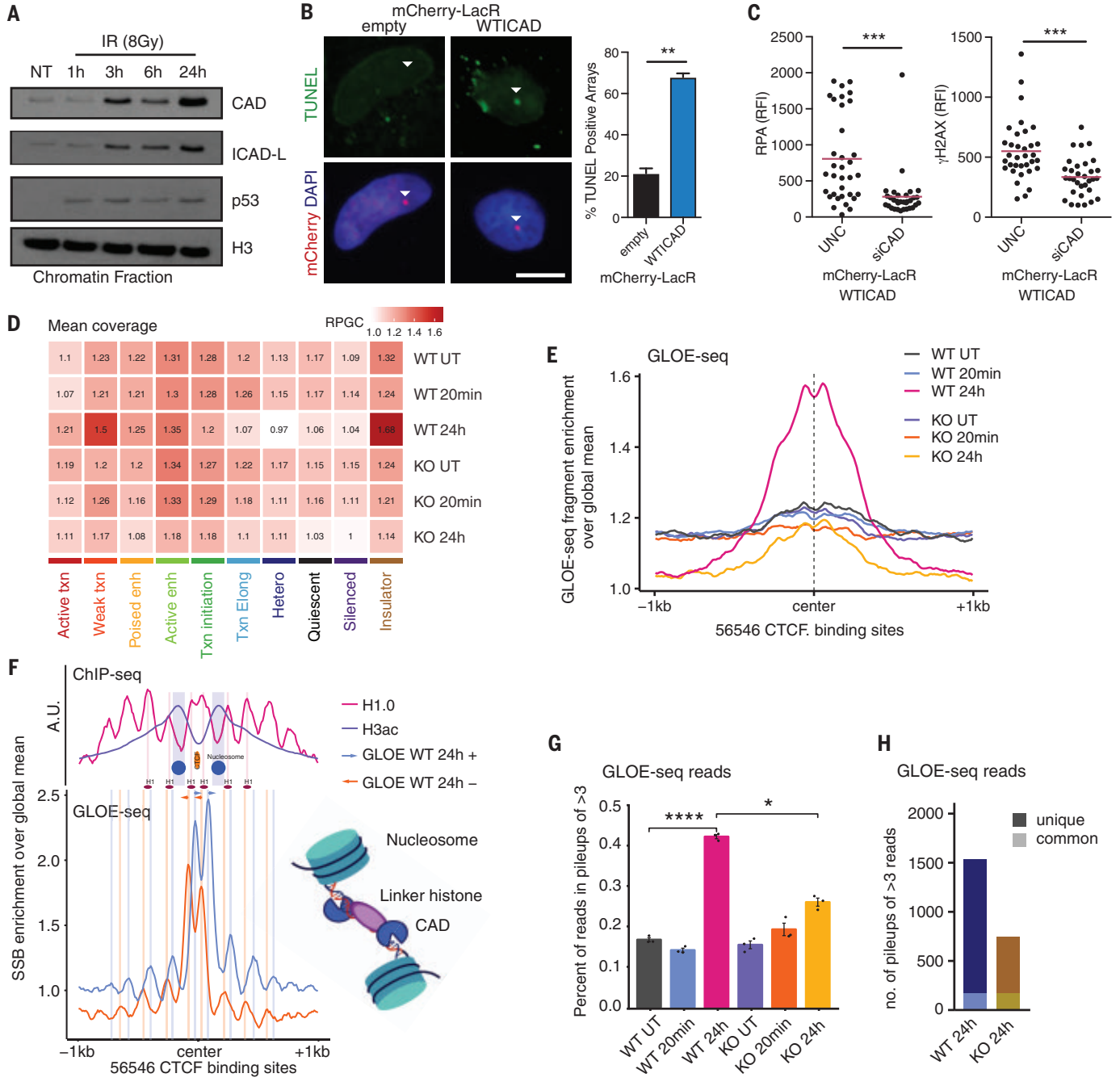


Fig. 2. CAD/ICAD chromatin recruitment inflicts DNA breaks at defined genomic elements. (**A**) Immunoblotting of chromatin fraction of HCT116 cells upon 8 Gy of IR. Recruitment of p53 was used as a positive control; H3 was used as a loading control. (**B**) TUNEL end labeling of 3′-OH indicates the formation of DNA breaks in mCherry-LacR-ICAD transfected cells. Scale bar, 10 μm. Data are means ± SEM; $N = 3$, $n > 50$. $^{**}P < 0.005$ (unpaired Student's t test). (**C**) Knockdown of CAD in mCherry-LacR-ICAD transfected cells reduces RPA and γ-H2AX. $N = 3$, $n > 20$. $^{***}P < 0.001$ (unpaired Student's t test). RFI, relative fluorescence intensity. (**D**) Genome-wide landscape of SSBs in HCT116 WT or CAD KO cells, before, 20 min after, and 24 hours after 8 Gy of IR. Average GLOE-seq read densities combined from three independent replicates are summarized over functionally distinct genomic regions as defined by a 10-state ChromHMM genome annotation. Elong, elongation; Enh, enhancer; Hetero, heterochromatin; Txn, transcription; UT, untreated. (**E**) SSB distribution around 56,546 CTCF binding sites in HCT116 cells. Average profiles from paired-end GLOE-seq fragments are plotted. (**F**) Footprint analysis of core and linker histones and SSBs around 56,546 CTCF binding sites in HCT116 cells. Linker histone H1.0 genomic occupancies are well positioned with a ~160-bp periodicity flanking the central CTCF binding site. GLOE-seq nick sites are piled up separately for the plus and minus strand. SSBs peak symmetrically and in a defined direction adjacent to the linker histone. (**G**) Three GLOE-seq replicates of each condition were filtered for duplicates and artificially over-represented regions, downsampled to the same sequencing depth (2 Mio reads), and examined for pileups of three or more unique forward reads within a 20-bp window indicating exact or near-exact coincidence of SSBs in multiple individual cells, thus "hypersensitive" nick sites. Data are means ± SEM. $^{*}P = 0.011$, $^{****}P = 5.6 \times 10^{-5}$ (Student's t test and Bonferroni correction). (**H**) A comprehensive list of pileups was generated from pooled replicates. Each pileup region was matched against the pileup regions of the other five conditions to exclude common treatment-independent hypersensitive nick sites. WT cells showed more than 1000 unique hypersensitive nick sites as compared to CAD KO cells.

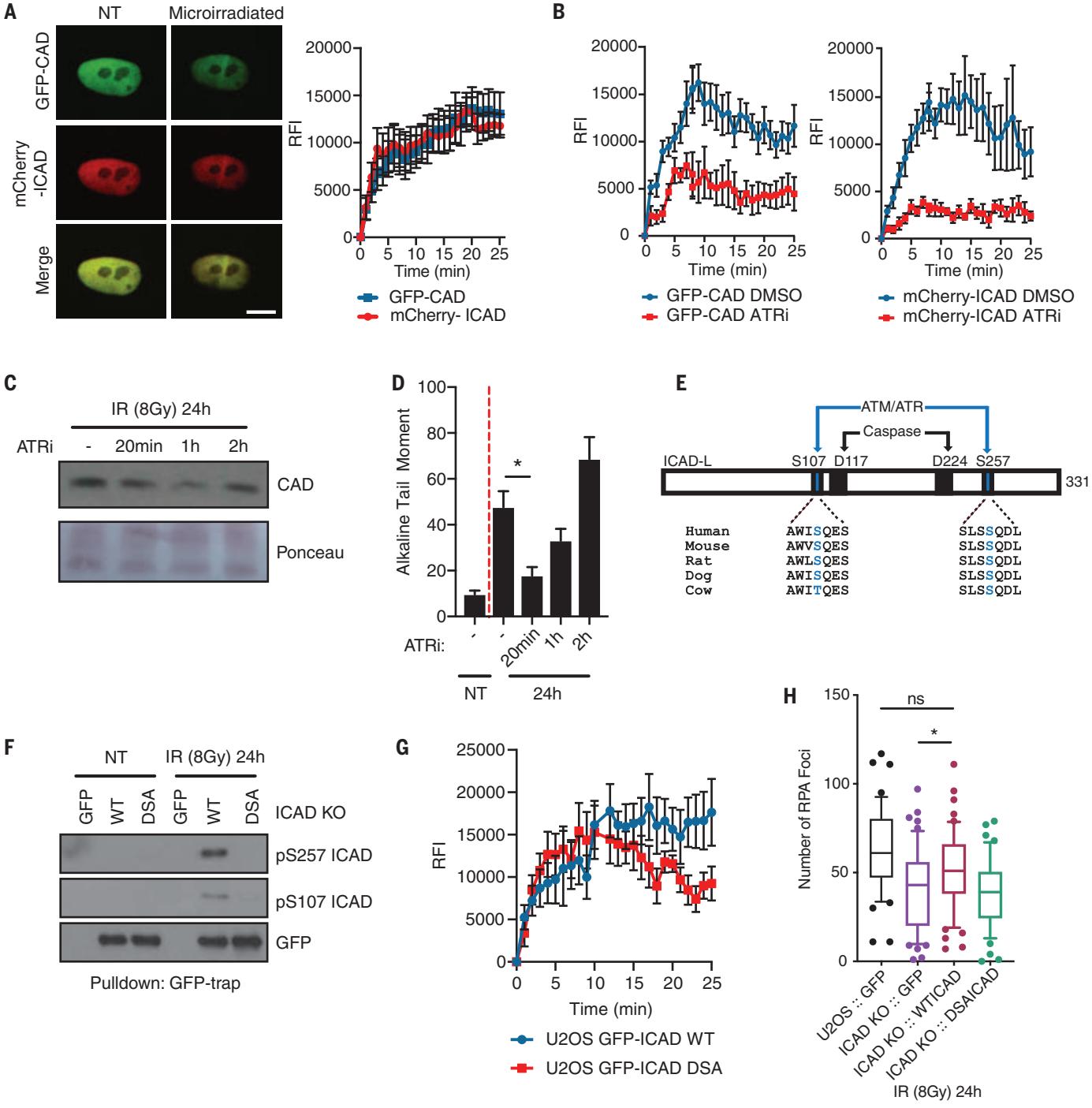


Fig. 3. ATM/ATR signaling to ICAD regulates CAD-induced DNA breaks after IR. (**A**) U2OS cells expressing mCherry-ICAD and GFP-CAD were microirradiated and imaged every minute for 25 min. Scale bar, 10 μm. Data are means ± SEM; cells per biological replicate, n = 6; biological replicates, N = 3. (**B**) Recruitment of CAD and ICAD to microirradiated laser stripes after inhibition of ATR. Conditions as in (A), except that ATRi AZ20 was added 20 min before microirradiation. RFI, SEM, replicates, and n/N are the same as in (A). (**C**) ATR inhibition (AZ20) for the indicated times prior to collection affects chromatin recruitment of CAD at 24 hours after 8 Gy of IR in HCT116 cells. (**D**) DNA break measurements under conditions like those in (C) in HCT116 cells. Data are means ± SEM; N = 3, n > 100. *P < 0.0443 (Kruskal-Wallis multiple-comparisons test). (**E**) Protein sequence analysis (Clustal Omega) of ICAD indicates two conserved ATM/ATR SQ phosphorylation motifs. (**F**) ICAD is phosphorylated at Ser^{107} and Ser^{257} after IR. ICAD KO U2OS cells expressing GFP, WT ICAD, or DSA ICAD were irradiated (8 Gy) and GFP-tagged proteins enriched by GFP trap. Immunoblotting with phospho-specific antibodies to Ser^{107} and Ser^{257} on ICAD. (**G**) DSA-ICAD is not stably recruited to laser-microirradiated damage. U2OS cells expressing GFP-WT ICAD or GFP-DSA ICAD were laser-microirradiated and imaged every minute for 25 min. Data are means ± SEM; cells per biological replicate, n = 7; biological replicates, N = 3. (**H**) RPA foci in indicated cells 24 hours after IR; N = 3, n > 50, box-and-whisker plot shows median and 10th to 90th percentiles. *P = 0.0324 [multiple-comparisons analysis of variance (ANOVA)].

(TUNEL) (Fig. 2B). Activation of the DNA damage response was noted by increased RPA association, ATR (ataxia telangiectasia and RAD3-related) association, and phosphorylation of histone variant H2AX (fig. S5, C to E). The DDR signaling from ICAD-tethered LacO arrays was dependent on CAD, as small interfering RNA (siRNA)–mediated knockdown of CAD reduced the recruitment of RPA and the phosphorylation of H2AX (Fig. 2C). Inhibition of caspase activity had no apparent effect on the recruitment of RPA to ICAD-tethered LacO arrays (fig. S5F). This demonstrates a capacity of CAD to mediate DNA break accumulation while associated with intact ICAD, independent of caspase signaling.

CAD-dependent DNA breaks at defined genomic loci

Together, our observations indicate that CAD-dependent DNA breaks appear to predominantly manifest as SSBs. To determine whether these breaks were occurring at defined genomic loci, we mapped SSBs at base-pair resolution in HCT116 wild-type or CAD KO cells before, 20 min after, and 24 hours after IR by GLOE-seq (genome-wide ligation of 3′-OH ends followed by sequencing, fig. S6A) (*24*). Examining natural SSB frequency as a function of different chromatin states, we found a weak prevalence of SSBs at active enhancers and transcription initiation sites, known to be associated with accessible chromatin, as well as insulators as defined by CTCF binding (Fig. 2, D and E). CTCF has multiple functions in genome biology as it assists the three-dimensional (3D) folding of chromatin by regulating the location of chromatin loops formation (*25*). This SSB distribution was also maintained immediately after irradiation and was independent of the presence of CAD enzyme (Fig. 2D). However, 24 hours after irradiation, a more distinctive distribution was observed in the presence of wild-type but not CAD KO cells: SSBs concentrated more on insulator regions while becoming relatively depleted in heterochromatic regions (Fig. 2, D and E). This shift in pattern was entirely consistent among replicates (fig. S6B). Consistent with the chromatin state analysis, we also observed a high enrichment of SSBs at and around CTCF binding sites 24 hours after irradiation, in a CAD-dependent manner (Fig. 2E).

We sought to more precisely pinpoint the SSBs around CTCF sites at base-pair resolution, using the mapping information of the first read, which identifies the exact nick ligation site and strand. Piling up nick sites around CTCF binding sites revealed a periodic pattern (Fig. 2F and fig. S7, A and B), with nicks in the plus and minus strand being separated by 185 base pairs. Nucleosomes are known to be well positioned around CTCF sites (*26*); hence, we used published acetylated histone H3 and linker histone chromatin immunoprecipitation sequencing (ChIP-seq) data to delineate the position of core and linker histones (*27*, *28*). Matching GLOE-seq with these ChIP-seq patterns revealed a consistent periodicity of core, linker histone, and SSB, with SSBs arising symmetrically left and right of the linker histone footprint (Fig. 2F). This suggested that DNA bound by the nucleosomes and linker histone DNA is protected from SSBs. The nick was introduced in a strand-specific manner, with the plus strand being nicked on the "plus" side of the linker histone, and the minus strand on the "minus" side of the linker histone; this suggests a potential role for histones (or chromatin structure in general) in directing and orienting CAD activity in a strand-specific manner.

CAD-induced SSBs appeared to be less random than naturally occurring SSBs. We noted initially that multiple unique GLOE-seq reads tended to accumulate at relatively few sites in the genome at 24 hours after irradiation, but not in any other condition. Thus, we used an unbiased approach to quantify SSB hotspots (more than three unique SSBs mapped in close proximity on the same strand). Twenty-four hours after irradiation in wild-type but not CAD KO cells, we observed a statistically significant ($P = 5.6 \times 10^{-5}$) factor of 2 increase across the three experimental replicates (Fig. 2G). A large number of the hypersensitivity sites that arose 24 hours after IR were new. We termed these CAD-dependent SSBs (CdSSBs), and they appeared to accumulate at hotspots different from those generally sensitive to occurrence of SSBs. Comparing wild-type and CAD KO cells 24 hours after irradiation, we observed 1371 unique pileups in wild-type cells, corresponding to putative CdSSBs, whereas ~581 unique pileups were present in the CAD KO but not wild-type cells (Fig. 2H). Notably, 232 of these 1371 CdSSBs overlapped with published CTCF peaks and 195 with DNase hypersensitive sites. Of the unique 581 pileups found in CAD KO cells, the overlap was only 11 and 14, respectively. In summary, genome-wide SSB mapping revealed a characteristic, unusual, CAD-dependent SSB landscape 24 hours after irradiation.

DDR signaling through ICAD coordinates CAD activity after genotoxic stress

Next, we investigated whether CAD and ICAD are embedded in the DDR signaling machinery in response to genotoxic stress. To explore this, we examined the recruitment of CAD/ICAD to stripes of microirradiated DNA in real time. Here, we observed that both ICAD and CAD were recruited with comparable kinetics to subnuclear regions of DNA breaks (Fig. 3A and movies S1 and S2). Unexpectedly, we noted that ICAD could also be recruited to chromatin after IR in the absence of CAD, which could indicate that the recruitment of CAD is mediated in part by ICAD after IR (fig. S8A). Examination of the recruitment to microirradiated DNA of a series of ICAD truncation fragments in ICAD-deficient cells indicated that all fragments could be recruited to irradiated regions. Fragments lacking the C-terminal domains of ICAD displayed a modest delay in accumulation (fig. S8, B to D). These results indicate that multiple regions of ICAD are responsible for chromatin recruitment.

Next, we examined whether a DDR signal could regulate the recruitment of CAD/ICAD to microirradiated stripes. We noted that inhibition or loss of ATM (ataxia telangiectasia mutated) and ATR kinase activity could limit the recruitment of CAD/ICAD (Fig. 3B and fig. S9, A to D). A similar role of the ATR kinase was also observed the day after IR, where ATR inhibition transiently diminished CAD chromatin recruitment as well as the corresponding number of DNA breaks (Fig. 3, C and D). To determine whether ATR regulates ICAD more directly, we analyzed the sequence of ICAD and identified two potential ATM/ATR phosphorylation sites, Ser^{107} and Ser^{257}, that are highly conserved across mammalian species (Fig. 3E). Further, the phosphorylation of Ser^{257} on ICAD was previously identified in phosphoproteome analysis of human cells exposed to IR (*29*). We assessed the propensity of ATR to phosphorylate ICAD in vitro, and demonstrated that ICAD can be directly phosphorylated by ATR, dependent on the ICAD SQ sites (fig. S9E). To further address the biological relevance of ICAD phosphorylation, we generated phospho-specific antibodies toward the Ser^{107} and Ser^{257} sites. Both sites appeared extensively phosphorylated after cell exposure to IR, and the phosphorylation was dependent on both ATM and ATR kinase activity 24 hours after IR (Fig. 3F and fig. S9, F to I). Next, we expressed a serine-to-alanine mutant form of ICAD that could not be phosphorylated (S107A and S257A; DSA) to investigate the functional relevance of these phosphorylation events (Fig. 3F and fig. S9F). Measuring recruitment to microirradiated stripes of DNA damage demonstrated that the DSA variant could not be stably recruited to these sites, unlike the wild-type ICAD (Fig. 3G). In addition, the DSA variant could not completely restore RPA foci formation at 24 hours after IR even though expression of CAD was restored (Fig. 3H and fig. S9F). This indicates that the ATR/ATM-dependent phosphorylation of ICAD functionally contributes to regulating the induction of DNA breaks and continued control of the checkpoint through CAD/ICAD after IR.

CAD is required for cell cycle checkpoint control and tumor cell survival after IR

Next, we investigated the role of CAD/ICAD-dependent maintenance of the G_2 cell cycle

checkpoint. We confirmed premature mitotic entry observed in the siRNA-based screen using individual siRNAs targeting CAD and ICAD and in the CAD-deficient KO cells (Fig. 4, A to C). This revealed that the breakdown in G_2 checkpoint control in CAD-deficient cells was most pronounced 24 hours after IR, which corresponded with our observed peak of CAD-inflicted DNA breaks (Fig. 4, B and C). To further characterize the molecular basis of the CAD/ICAD-dependent checkpoint regulation, we noted reduction of the inhibitory phosphorylation of cyclin-dependent kinase CDK1 (Tyr[15]), reduction in active checkpoint kinase CHK2, and moderate difference in

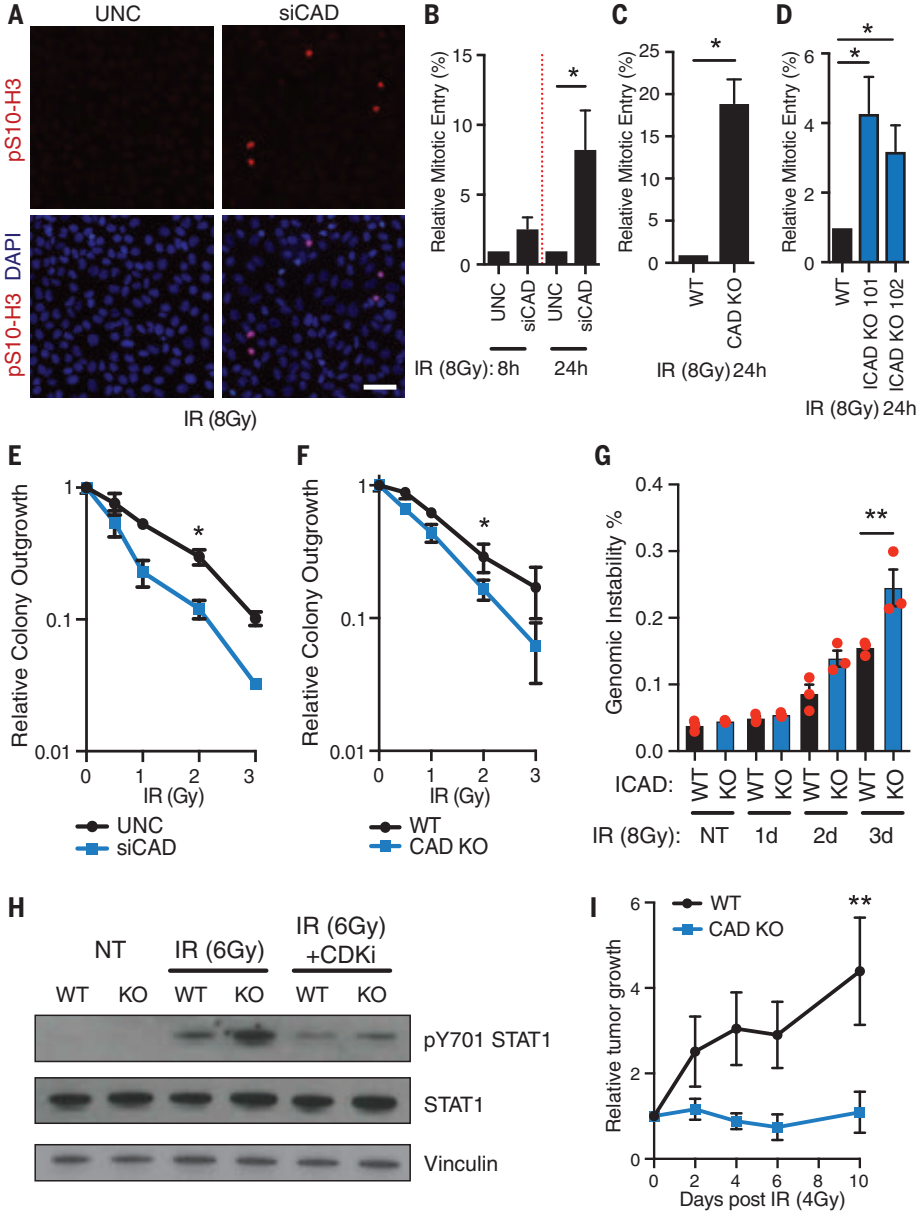


Fig. 4. CAD/ICAD are required for tumor cell survival and checkpoint maintenance after IR. (**A**) Mitotic (pS10-H3) staining of irradiated and nocodazole (NZ)–trapped U2OS cells transfected with control siRNA (UNC) and siRNA against CAD. Cells were irradiated with 6 Gy of IR; after a 2-hour recovery, NZ was added for 8 hours before cells were fixed for immunofluorescence. Scale bar, 50 μm. (**B**) G_2 checkpoint maintenance in HCT116 CAD siRNA knockdown cells upon 8 Gy of IR. Data are means ± SEM; $N = 3$. $^*P < 0.05$ (unpaired Student's t test). (**C**) G_2 checkpoint maintenance in HCT116 CAD KO cells 24 hours after 8 Gy of IR. Data are means ± SEM; $N = 3$. $^*P < 0.05$ (unpaired Student's t test). (**D**) G_2 checkpoint maintenance in U2OS ICAD KO cells 24 hours after 6 Gy of IR. Data are means ± SEM; $N = 3$. $^*P < 0.05$ (unpaired Student's t test). (**E**) Relative colony outgrowth of CAD-depleted cells after IR exposures. Data are means ± SEM; $N = 4$. $^*P < 0.05$ (unpaired Student's t test). (**F**) Relative colony outgrowth of CAD KO HCT116 cells upon indicated IR exposure. Data are means ± SEM; $N = 4$. $^*P < 0.05$ (unpaired Student's t test). (**G**) Genomic instability in U2OS WT and ICAD KO cells upon 8 Gy of IR. % Genomic instability represents total number of cells displaying micronuclei and fragmented nuclei divided by the total number of cells. Data are means ± SEM; $N = 3$, $n > 400$. $^{**}P = 0.0016$ (ANOVA). (**H**) Immunoblotting of pY701 STAT1 3 days after 6 Gy of IR in WT and ICAD KO cells. CDK1 inhibitor RO-3306 was added 2 hours after IR. (**I**) Normalized tumor growth of HCT116 WT and HCT116 CAD KO tumors after 4 Gy of IR. Data are means ± SEM; $n = 6$. $^{**}P = 0.0055$ (two-way multiple-comparisons ANOVA). (**J**) Model of CAD-dependent G_2 phase checkpoint.

active CHK1 (fig. S10, A to E). CDK1 inhibition is required to restrict mitotic entry after IR. Additionally, activated CHK2 and the phosphorylation of KAP1 have also been implicated in controlling mitotic entry of cells after IR (*30*). In line with our previous observations, this checkpoint function appeared independent from caspase signaling and unrelated to apoptotic cell death (fig. S11, A to C). Further, a CAD-promoted G_2 checkpoint was not observed in nonmalignant cells (fig. S11D), as normal cells harbor proficient p53 and pRB pathways that ensure robust G_1/S transition control and diminished G_2 checkpoint dependency. The timing of G_2 checkpoint breakdown in CAD-deficient cancer cells corresponded to the kinetics of CAD-inflicted DNA breaks, indicating that this DNA modification may functionally prevent premature mitotic entry after IR. Notably, CAD/ICAD-deficient cells that entered mitosis prematurely exhibited a high number of lagging chromosomes and chromatin bridges (fig. S11E).

Aberrant cell cycle progression and premature mitotic entry with unrepaired DSBs leads to mitotic cell death, which contributes to radiosensitivity (*21*, *31*). Indeed, cancer cells lacking the expression of CAD or ICAD displayed increased radiosensitivity (Fig. 4, E and F, and fig. S12A). This phenomenon was selective for cancer cells, as loss of CAD in nonmalignant cells had no impact (fig. S12B). Failure to repair DNA breaks through the inhibition of PARP activity has been suggested to sensitize cells to IR (*32*). Therefore, we examined whether failure to repair CAD-inflicted DNA breaks would contribute to this sensitization. We found that CAD-proficient cancer cells were sensitized to radiation by the addition of PARP inhibitor 24 hours after IR. However, PARP-inhibited CAD-deficient cells did not display any additional sensitization to IR (fig. S12C). Additionally, we noted that the loss of G_2 cell cycle checkpoint control led to the increased incidence of unstable nuclei (micronuclei and fragmented nuclei) (Fig. 4G and fig. S11E). Such genomic instability is considered to be a potent molecular pattern signal that activates inflammatory STAT1 signaling (*33*). Thus, we assessed the activating phosphorylation of STAT1 at Tyr^{701} after IR in wild-type and CAD-deficient cells. This revealed elevated p-Tyr^{701} STAT1 in CAD-deficient cells after IR, which was dependent on progression through mitosis (Fig. 4H). These observations are in line with recent research demonstrating that premature progression through mitosis after radiation promotes STAT1 signaling, and our data further indicate that CAD/ICAD limits this response (*33*, *34*).

To complement the cell-based observations, we analyzed CAD function in a model of tumor radiotherapy in vivo, using human tumor xenografts and tumor growth after radiation (Fig. 4I and fig. S12, D to F). Consistent with the cell-based survival assessment, a pronounced negative impact on tumor growth was detected after irradiation in tumors deficient in CAD, relative to their CAD-proficient counterparts. The irradiated CAD-deficient tumors demonstrated elevated p-Tyr^{701} STAT1 compared to the wild-type tumors at endpoint (fig. S12G). Collectively, these results support the concept that the CAD-dependent pathway actively promotes cancer cell survival after IR. Previous murine studies had implicated CAD as a potential tumor suppressor, which was linked to pro-apoptotic function (*10*). However, analysis of gene expression data comparing normal and malignant tissues in human cancers indicated that loss of function of CAD/ICAD is a rare event (fig. S13, A and B) (*35*). Further, elevated expression of CAD in particular was noted in multiple tumor types (fig. S13, A and B)—an observation that is consistent with a potential, as yet unidentified, tumor-supporting role for CAD.

Discussion

Together, our results unravel a DDR-mediated G_2 phase checkpoint pathway where cancer cells exposed to IR inflict reversible CAD-dependent DNA breaks including the CdSSBs. These lesions stimulate signaling responses and prevent premature mitotic entry (Fig. 4J), which enhances cancer cell survival. As repair of IR-induced DNA damage progresses, the number of highly genotoxic complex DNA DSBs declines, dropping below a threshold required to maintain the checkpoint. In turn, the induction of CAD-dependent DNA breaks signals an amplification of the DDR, further stabilizing the G_2 checkpoint and thereby providing more time for repair of the more complex, difficult-to-repair, and potentially lethal IR-induced genotoxic lesions.

The observations presented here indicate that the CAD-mediated checkpoint signal is primarily dependent on the generation of CdSSBs, a form of DNA damage characterized by rapid repair kinetics. Further, we implicate the activity of PARP-1 in the repair of CAD-mediated DNA breaks, as the addition of a nicotinamide adenine dinucleotide (NAD)–like PARP-1 inhibitor, 4-ANI, promotes DNA break formation in a CAD-dependent manner. Inhibition of PARP-1 activity has been reported to impair the G_1 checkpoint but enhance the G_2 checkpoint in irradiated cells (*36*), which may be the result of loss of PARP-1–directed repair of CdSSBs leading to increased DNA lesion burden. Furthermore, we uncovered an activating role of ICAD phosphorylation mediated by the DNA damage–induced ATM and ATR kinases. This may suggest an ongoing DNA damage–mediated feedback loop that is active until repair of complex lesions is completed, at which stage the kinase signaling declines. In this regard, we note that multiple posttranslational modification sites have been identified in ICAD (www.phosphosite.org/proteinAction.action?id=9541&showAllSites=true). These additional sites suggest that multiple signaling events may converge to regulate the CAD/ICAD checkpoint pathway.

Prior observations had established a temporally delayed secondary wave of SSBs after IR; however, the origin and functional role of these lesions have remained obscure. We identified CdSSBs as the source of these lesions, which accumulate at a subset of CTCF sites in the genome. These genomic loci may be accompanied by a moderate number of CAD-dependent stochastic lesions that escape detection because of rare targeting events within unspecified regions. Functionally, CTCF sites serve as binding domains for the CTCF protein, which regulates 3D chromosomal looping and topologically associated domains (TADs) in interphase cells (*25*). In response to positioned DSBs, chromosomal loops form in proximity to the CTCF sites, which act to sculpt the chromatin spreading of the phosphorylated histone variant γH2AX. Given that CdSSBs extend cell cycle checkpoint control, an appealing concept is that the newly formed SSBs help to enforce a chromatin response to DNA damage. This is supported by the CAD-dependent phosphorylation of KAP1, a major chromatin marker of ongoing DDR. Here, CAD nuclease activity is restricted to generate CTCF-directed SSBs after IR, which is in contrast to CAD-dependent genome-wide intranucleosomal cleavage events during apoptosis. Hence, the precision of strand break formation may be a determinant in how CAD guides cancer cell survival.

On the basis of our present study, we propose that CAD/ICAD signaling is an adaptive cancer-intrinsic mechanism to resist genotoxic stress. The apparent selectivity of this stress-tolerance pathway likely reflects multiple factors that occur in cancer but not normal cells, including defects in the p53 and pRB pathways controlling G_1/S transition, oncogene-driven premature S-phase entry, enhanced replication stress, and defective DNA repair, as well as checkpoint signaling mechanisms. The CAD/ICAD-mediated pathway may reflect adaptation to the genome-destabilizing selective pressures during tumorigenesis and contribute to therapy resistance. Such a prosurvival checkpoint pathway also reveals a cancer-selective vulnerability, thereby providing a potential avenue to enhance tumor cell radiosensitivity by targeting this G_2 cell cycle checkpoint.

REFERENCES AND NOTES

1. R. A. Sharma *et al.*, *Nat. Rev. Clin. Oncol.* **13**, 627–642 (2016).
2. T. D. Halazonetis, V. G. Gorgoulis, J. Bartek, *Science* **319**, 1352–1355 (2008).

3. M. T. Dillon, J. S. Good, K. J. Harrington, *Clin. Oncol.* **26**, 257–265 (2014).
4. A. Cieślar-Pobuda, Y. Saenko, J. Rzeszowska-Wolny, *Mutat. Res.* **732**, 9–15 (2012).
5. D. K. Klein *et al.*, *Nat. Commun.* **6**, 5800 (2015).
6. A. N. Kousholt *et al.*, *J. Cell Biol.* **197**, 869–876 (2012).
7. M. Enari *et al.*, *Nature* **391**, 43–50 (1998).
8. K. Samejima, W. C. Earnshaw, *Nat. Rev. Mol. Cell Biol.* **6**, 677–688 (2005).
9. R. A. V. Bell, L. A. Megeney, *Cell Death Differ.* **24**, 1359–1368 (2017).
10. B. D. Larsen, C. S. Sørensen, *FEBS J.* **284**, 1160–1170 (2017).
11. J. H. Song, K. Kandasamy, M. Zemskova, Y. W. Lin, A. S. Kraft, *Cancer Res.* **71**, 506–515 (2011).
12. J. D. Orth, A. Loewer, G. Lahav, T. J. Mitchison, *Mol. Biol. Cell* **23**, 567–576 (2012).
13. X. Liu *et al.*, *Cell Res.* **27**, 764–783 (2017).
14. G. Ichim *et al.*, *Mol. Cell* **57**, 860–872 (2015).
15. B. D. Larsen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 4230–4235 (2010).
16. J. S. Brown, B. O'Carrigan, S. P. Jackson, T. A. Yap, *Cancer Discov.* **7**, 20–37 (2017).
17. J. Murai *et al.*, *Cancer Res.* **72**, 5588–5599 (2012).
18. X. Li *et al.*, *J. Biol. Chem.* **282**, 36177–36189 (2007).
19. M. H. Al-Khalaf *et al.*, *Cell Discov.* **2**, 15041 (2016).
20. V. Iglesias-Guimarais *et al.*, *J. Biol. Chem.* **288**, 9200–9215 (2013).
21. R. G. Syljuåsen *et al.*, *Cancer Res.* **64**, 9035–9040 (2004).
22. H. Sakahira, Y. Takemura, S. Nagata, *Arch. Biochem. Biophys.* **388**, 91–99 (2001).
23. S. M. Janicki *et al.*, *Cell* **116**, 683–698 (2004).
24. A. M. Sriramachandran *et al.*, *Mol. Cell* **78**, 975–985.e7 (2020).
25. C. Arnould, G. Legube, *J. Mol. Biol.* **432**, 724–736 (2020).
26. Y. Fu, M. Sinha, C. L. Peterson, Z. Weng, *PLOS Genet.* **4**, e1000138 (2008).
27. S. Walz *et al.*, *Nature* **511**, 483–487 (2014).
28. C. M. Torres *et al.*, *Science* **353**, aaf1644 (2016).
29. S. Matsuoka *et al.*, *Science* **316**, 1160–1166 (2007).
30. H. Jaiswal *et al.*, *EMBO J.* **36**, 2161–2176 (2017).
31. L. Galluzzi *et al.*, *Cell Death Differ.* **25**, 486–541 (2018).
32. A. Dréan, C. J. Lord, A. Ashworth, *Crit. Rev. Oncol. Hematol.* **108**, 73–85 (2016).
33. S. M. Harding *et al.*, *Nature* **548**, 466–470 (2017).
34. J. Chen *et al.*, *Cell Rep.* **32**, 108080 (2020).
35. A. R. Cortazar *et al.*, *Cancer Res.* **78**, 6320–6328 (2018).
36. M. Masutani, T. Nozaki, K. Wakabayashi, T. Sugimura, *Biochimie* **77**, 462–465 (1995).

ACKNOWLEDGMENTS

We thank the core facilities at BRIC for assistance; the CSS laboratory for insightful comments; A. H. Lund (Biotech Research and Innovation Centre, University of Copenhagen) for HCT116 p53 KO cells; D. Spector (Cold Spring Harbor Laboratory) for U2OS 263 cells; D. Durocher (Lunenfeld-Tanenbaum Research Institute) for mCherry-LacR expression plasmid; F. Zhang (McGovern Institute, Massachusetts Institute of Technology) for pSpCas9(BB)-2A-GFP and pSpCas9(BB)-2A-GFP plasmids; W. Earnshaw (Wellcome Centre for Cell Biology, University of Edinburgh) for murine CAD/ICAD coexpression plasmid (pRHS); and B. Singers Sørensen (University of Aarhus) for discussions and cell lines used during revision. **Funding:** Supported by Danish Cancer Society grants R90-A5949 (C.S.S.), R204-A12415 (J.Be.), and R204-A12617-B153 (J.Ba.); Danish Council for Independent Research grant 4004-00621 (C.S.S.); Novo Nordisk Foundation grants NNF16OC0022358 (C.S.S.) and NNF 0060590 (J.Ba.); Canadian Institute of Health Research grant 156120 (L.A.M.); a postdoctoral grant from Independent Research Fund Denmark (B.D.L.); Czech Science Foundation grant 19-07674S and Swiss National Science Foundation grant 310030_184716 (P.J.); Swedish Research council grant VR-MH 2014-46602-117891-30, Danish National Research Foundation (project CARD) grant DNRF 125, and Grant Agency of the Czech Ministry of Health grant NU21-03-00195 (J.Ba.); The European Union's Horizon 2020 program under the Marie Sklodowska-Curie grant (agreement 722729) (G.P.); and the Karolinska Institutet SFO for Molecular Biosciences, Vetenskapsrådet Junior Researcher grant 2015-04815, and H2020 ERC Starting Grant 715024 RAPID (S.J.E.). Bioinformatics analyses were performed on resources provided by the Swedish National Infrastructure for Computing (SNIC) at Uppmax server (projects SNIC 2020/15-9, SNIC 2020/6-3 to S.J.E.). **Author contributions:** Conceptualization: B.D.L., C.S.S. Methodology: B.D.L., J.Be., P.Y.K.Y., R.B., G.P., S.J.E., C.S.S. Experimental work: B.D.L., J.Be., P.Y.K.Y., R.B., G.P., V.U., J.K.A., T.T.K., S.E. Data documentation: B.D.L., J.Be., R.B., G.P., S.J.E., C.S.S. Data interpretation: B.D.L., J.Be., P.J., L.A.M., S.J.E., J.Ba., C.S.S. Funding acquisition: B.D.L., J.Be., P.J., L.A.M., S.J.E., J.Ba., C.S.S. Project administration: B.D.L., C.S.S. Supervision: B.D.L., P.J., L.A.M., S.J.E., J.Ba., C.S.S. Writing–original draft: B.D.L., J.Be., C.S.S. Writing–review and editing: B.D.L., J.Be., P.Y.K.Y., R.B., L.A.M., S.J.E., J.Ba., C.S.S. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All materials are available upon request from C.S.S. All sequencing data and code are available at www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE171242 and https://doi.org/10.5281/zenodo.6386176.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abi6378
Materials and Methods
Figs. S1 to S13
Tables S1 and S2
Movies S1 and S2
References (*37*–*44*)
MDAR Reproducibility Checklist

30 March 2021; resubmitted 4 January 2022
Accepted 30 March 2022
10.1126/science.abi6378

NEUROSCIENCE

Molecular and neural basis of pleasant touch sensation

Benlong Liu[1]†, Lina Qiao[1]†‡, Kun Liu[1]†§, Juan Liu[1], Tyler J. Piccinni-Ash[1], Zhou-Feng Chen[1,2]*

Pleasant touch provides emotional and psychological support that helps mitigate social isolation and stress. However, the underlying mechanisms remain poorly understood. Using a pleasant touch–conditioned place preference (PT-CPP) test, we show that genetic ablation of spinal excitatory interneurons expressing prokineticin receptor 2 (PROKR2), or its ligand PROK2 in sensory neurons, abolishes PT-CPP without impairing pain and itch behaviors in mice. Mutant mice display profound impairments in stress response and prosocial behaviors. Moreover, PROKR2 neurons respond most vigorously to gentle stroking and encode reward value. Collectively, we identify PROK2 as a long-sought neuropeptide that encodes and transmits pleasant touch to spinal PROKR2 neurons. These findings may have important implications for elucidating mechanisms by which pleasant touch deprivation contributes to social avoidance behavior and mental disorders.

Our sense of touch is composed of discriminative and affective components. Discriminative touch detects physical properties of tactile stimuli (e.g., location, shape, texture, force, etc.), whereas affective touch conveys emotional value that is modulated by social context (*1*, *2*). Pleasant touch (e.g., cuddling, caressing, and hugging) encodes positive hedonic information that facilitates emotional development, affiliative behavior, and the well-being of social animals (*1*, *3*, *4*). Social touch is one of the most favored activities that might be evolutionarily conserved throughout the animal kingdom (*5*, *6*). In nonhuman primates, rodents, birds, and insects, allogrooming behavior (or allopreening for birds) is important for strengthening and maintaining social bonding, reciprocity, attachment, and hierarchy (*7*–*9*). Acute social isolation increases social cravings and reward-seeking behavior (*10*). Harlow's pioneering work demonstrated that infant rhesus monkeys separated from their mothers have an innate desire to cuddle soft cloth for contact comfort and emotional needs, and maternal touch is vital for the behavioral and psychological development of offspring (*11*). Likewise, long-term deprivation of maternal care and positive social touch has lasting negative consequences on the mental health of children (*12*, *13*). In fact, affective touch avoidance and deficiency are some of the hallmarks of several neuropsychiatric disorders, including autism spectrum disorders (ASDs) (*14*, *15*). Studies in humans have shown that C tactile (CT) fibers innervating hairy skin encode positive valence of social touch (*1*, *16*–*19*), whereas *MrgprB4*-expressing sensory neurons and *Gpr83*-expressing spinal projection neurons have been implicated in mice (*20*, *21*). Despite its profound importance, how pleasant touch information is encoded and transmitted from somatosensory neurons to the spinal cord remains unknown. Our understanding of the molecules and neural circuits of pleasant touch has been hampered by a paucity of suitable animal models and methodologies that permit accurate inference and assessment of the affective state of mice that experience pleasant touch. Unlike discriminative touch, affective touch mediated by unmyelinated C fibers is a slow process (*1*). We postulated that pleasant touch

[1]Center for the Study of Itch and Sensory Disorders and Department of Anesthesiology, Washington University School of Medicine, St. Louis, MO 63110, USA. [2]Departments of Medicine, Psychiatry, and Developmental Biology, Washington University School of Medicine, St. Louis, MO 63110, USA.
***Corresponding author. Email: chenz@wustl.edu**
†These authors contributed equally to this work.
‡Present address: Department of Biochemistry and Molecular Biology, Institute of Acupuncture and Moxibustion, China Academy of Chinese Medical Sciences, Beijing 100700, China.
§Present address: Department of Physiology, Institute of Acupuncture and Moxibustion, China Academy of Chinese Medical Sciences, Beijing 100700, China.

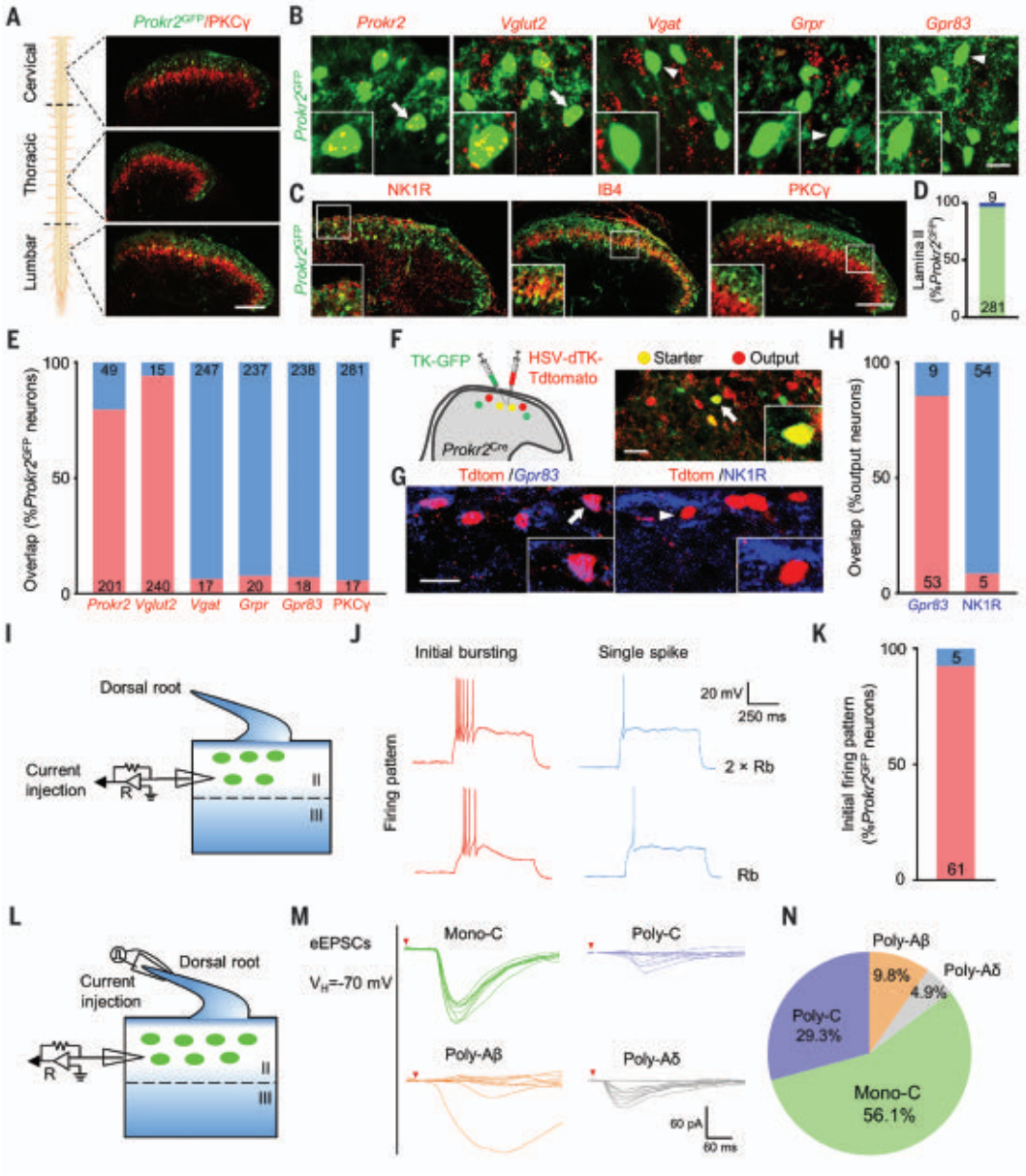


Fig. 1. PROKR2 neurons are a unique population of spinal excitatory interneurons. (**A** to **E**) Double staining of GFP with PKCγ at three levels of the spinal cord (A) and GFP with various markers in the lumbar cord [(B) and (C)]. Arrows indicate double-labeled cells, and arrowheads indicate GFP only. *n* = 3 mice. Scale bars, 100 μm [(A) and (C)] and 20 μm (B). (D) Quantification of (C), percentage of *Prokr2*^{GFP} cells in lamina II (green) and lamina I (blue). (E) Quantification of (B) and (C). Red in (E) indicates the percentage of double-labeled cells of *Prokr2*^{GFP} cells, and blue indicates GFP only. (**F**) Schematic of intraspinal injection of virus AAV5-Ef1a-DIO-EGFP-2a-TK-WPRE-pA (TK-GFP) and HSV-dTK-LSL-tdTomato (HSV-dTK-Tdtomato) in the dorsal horn of *Prokr2*^{Cre} mice at the lumbar level (left). (Right) Image showing virus expression in a starter neuron (yellow) expressing GFP and Tdtomato. Scale bar, 20 μm. (**G**) Double staining of *Gpr83* or NK1R anterogradely labeled with Tdtomato. Scale bar, 20 μm. (**H**) Quantification of (G). *n* = 3. (**I** to **K**) Schematic of the whole-cell patch-clamp recording of *Prokr2*^{GFP} neurons in the spinal cord slice preparations (I), a representative trace of the initial firing pattern (red) and single-spike firing pattern (blue) at 20 pA [rheobase (Rb)] and 40 pA [twofold rheobase (2 × Rb)] (J), and proportions of different types of firing pattern (K). *n* = 66 neurons. (**L** to **N**) Schematic of the recording of the type of inputs onto *Prokr2*^{GFP} neurons with dorsal root stimulation (L), representative traces of different types of inputs (M), and their proportions (N). eEPSCs, evoked excitatory postsynaptic currents. *n* = 41 neurons. All data are presented as means ± SEMs.

is encoded by slow-acting neuropeptides in C fibers and their cognate excitatory G protein–coupled receptors (GPCRs) in laminae II of the spinal cord that consist of microcircuits for relaying discrete sensory modalities from primary afferents to the brain (*22*). In a search for lamina- and modality-specific GPCRs, we found that prokineticin receptor 2 (PROKR2) is uniquely expressed in laminae II of the spinal cord, and PROKR2 neurons represent a previously unknown population of spinal excitatory interneurons. Using an unbiased behavioral paradigm in combination with physiological tests, extracellular recording, and genetic approaches, we sought to examine the role of the PROK2-PROKR2 signaling in pleasant touch.

Properties of spinal PROKR2-expressing lamina II excitatory interneurons

We used *Prokr2*^{GFP} transgenic mice as a surrogate to characterize PROKR2 expression in the spinal cord. RNAscope in situ hybridization (ISH) followed by immunohistochemical (IHC) studies indicated that green fluorescent protein (GFP) of *Prokr2*^{GFP} mice recapitulates a large part of *Prokr2* expression in laminae II (80.4%) of the spinal cord (Fig. 1, A to E). A significant fraction of GFP is distributed in the dorsal side of the lamina II inner (IIi) layer innervated by isolectin B4 (IB4)–binding non-peptidergic fibers (Fig. 1C). *Prokr2* is largely colocalized with the excitatory genes, like vesicular glutamate transporter 2 (*Vglut2*) (94.1%) and *Lmx1b* (96.8%), but rarely overlaps with the inhibitory neuronal markers, such as vesicular γ-aminobutyric acid (GABA) transporter (*Vgat*) (6.4%) or Pax2 (3.1%) (Fig. 1, B and E, and fig. S1, A and B). RNAscope ISH and IHC showed that *Prokr2* rarely overlaps with gastrin-releasing peptide receptor (*Grpr*) (7.8%), an itch-specific receptor expressed in laminae I and II interneurons (*23*, *24*), or protein kinase C γ (PKCγ) (5.7%), which labels the ventral side of the lamina IIi layer, with *Gpr83* (7.0%) or

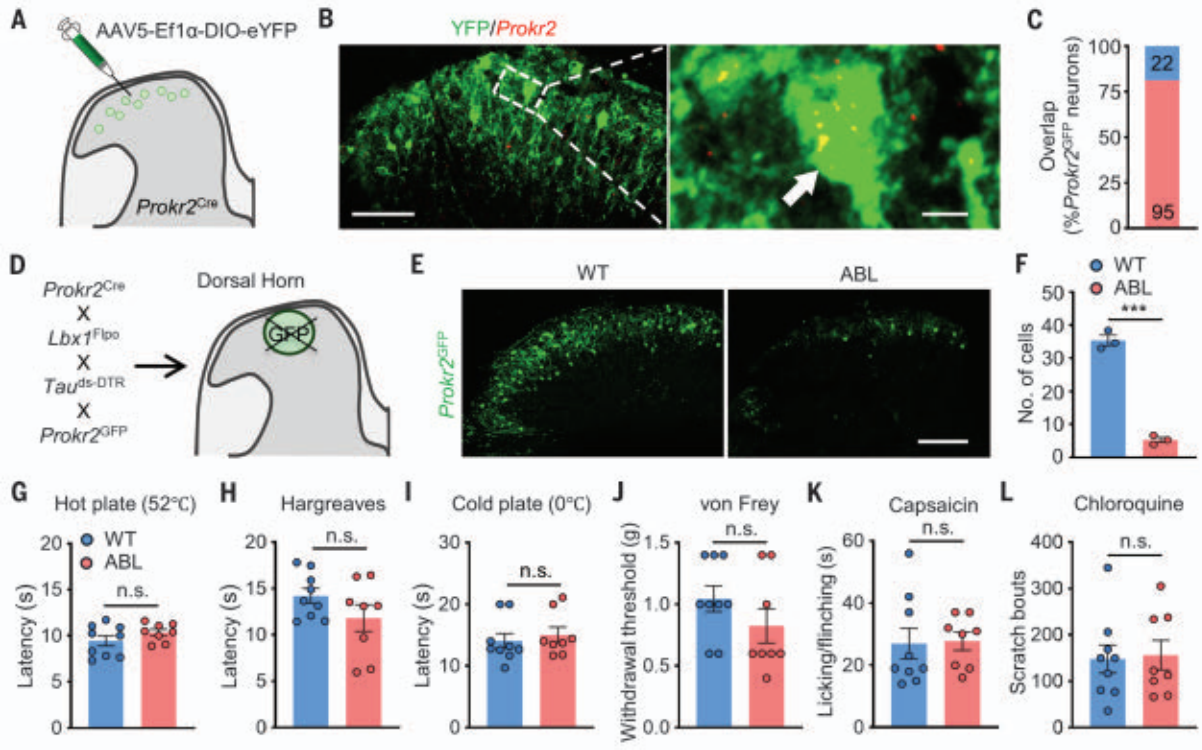


Fig. 2. PROKR2 neurons are dispensable for acute pain and itch behaviors. (**A**) Schematic of intraspinal injection of Cre-dependent adeno-associated virus (AAV) virus expressing YFP. (**B** and **C**) Double staining of YFP and *Prokr2* (arrow indicates a double-stained cell) in the lumbar cord (B) and percentage of overlapping cells (C). Scale bars, 50 μm (left) and 5 μm (right). (**D**) Strategy for intersectional genetic ablation of spinal *Prokr2*^{Cre} neurons. *Prokr2*^{Cre} mice were mated with *Lbx1*^{Flpo}, *Tau*$^{ds\text{-}DTR}$, and *Prokr2*^{GFP} lines to generate diphtheria toxin receptor (DTR)–expressing *Prokr2* neurons by injection of diphtheria toxin (50 μg/kg, i.p.). (**E** and **F**) Images of *Prokr2*^{GFP} neurons in WT and ABL mice (E) and quantification of *Prokr2*^{GFP} neurons (F). Scale bar, 150 μm. (**G** to **L**) Comparable latencies in hot plate (G), Hargreaves (H), and cold plate (I) tests; withdraw threshold in von Frey test (J); licking or flinching time induced by capsaicin (2 g, i.pl.) (K); and scratching numbers induced by chloroquine (200 μg, i.d.) (L) between WT and ABL mice. i.pl., intraplantar injection; i.p., intraperitoneal injection; i.d., intradermal injection. $n = 3$ (F); $n = 8$ to 9 [(G) to (L)]. Statistics by unpaired t test in (F) and (G) to (I). ***$P < 0.001$; n.s., not significant. Error bars indicate SEMs.

NK1R—two markers expressed in distinct subsets of projection neurons—implicated in affective touch and pain, respectively (*20*) (Fig. 1, B, C, and E). To identify the downstream target of PROKR2 neurons, we next performed Cre-dependent virus-mediated anterograde monosynaptic tracing in the spinal cord of *Prokr2*^{Cre} mice using HSV-dTK-LSL-tdTomato as the monosynaptic tracer to follow the output neurons (Fig. 1F). RNAscope ISH or IHC revealed that ~85.5% of the output neurons labeled by tdTomato express *Gpr83*, whereas only 8.5% express NK1R (Fig. 1, G and H).

We next examined whole-cell patch-clamp recordings of spinal cord slices obtained from *Prokr2*^{GFP} mice. The firing pattern of most *Prokr2*^{GFP} neurons was very homogeneous: ~92.4% displayed initial bursting firing, with a few neurons showing single-spike firing (Fig. 1, I to K). To identify the type of peripheral sensory inputs, we recorded the response of *Prokr2*^{GFP} neurons in the parasagittal section of spinal slices obtained from *Prokr2*^{GFP} mice with the dorsal root (L4 to L5) attached, using the dorsal root stimulation method as described previously (Fig. 1L) (*25*). *Prokr2*^{GFP} neurons predominantly receive monosynaptic and polysynaptic C fiber inputs (a combined 85.4%) with a small fraction of polysynaptic Aδ (4.9%) or Aβ inputs (9.8%) (Fig. 1, M and N). Furthermore, PROK2 application evoked subthreshold depolarizations in most *Prokr2*^{GFP} neurons, which were blocked by PKRA7, a PROKR2 antagonist (fig. S1, C and D).

PROKR2 neurons are dispensable for pain and itch transmission

To assess the function of spinal PROKR2 neurons, we first validated *Prokr2*^{Cre} mice by intraspinal injection of Cre-dependent yellow fluorescent protein (YFP) virus. Most *Prokr2*$^{Cre\text{-}GFP}$ neurons expressed *Prokr2* (81.2%) (Fig. 2, A to C). We then used an intersectional genetic strategy to ablate spinal PROKR2 neurons in *Prokr2*^{Cre} mice, hereafter referred to as ABL mice, as previously described (*25*) (Fig. 2, D to F, and fig. S2, A and B). The specificity of ablation was demonstrated by the loss of most *Prokr2*^{Cre} neurons (84.9%) and *Prokr2* (81.3%) without affecting PKCγ, neurokinin B (NKB), and *Grpr* in the spinal cord or *Prokr2*^{GFP} in discrete brain regions (Fig. 2, D to F; fig. S2, C and D; and fig. S3). A battery of pain and itch tests was conducted to assess the role of *Prokr2*^{Cre} neurons in somatosensory transmission. We did not observe a statistically significant difference in thermal pain, cold pain, mechanical thresholds, or inflammatory pain induced by capsaicin between ABL mice and their littermate control mice [referred to as wild-type (WT) mice; Fig. 2, G to K]. Chemical and mechanical itch as well as the hairy skin sensitivity to a sticky tape of ABL mice were comparable to those of WT mice (Fig. 2L and fig. S2, E and F).

PROKR2 neurons convey pleasant touch sensation

As a weak stimulus, gentle stroking with a soft brush on the hairy skin of mice does not elicit a robust motor response that can be used as a proxy to quantify pleasant touch sensation. Moreover, using a hand or soft brush to simulate pleasant touch in mice often causes avoidance behavior without habituation and conditioning. To circumvent these confounds, we developed a protocol that includes two procedural features to avoid the stress and

anxiety associated with handling. First, given that social isolation increases the urge for social touch and attachment (*26*), mice were single-housed for 1 week followed by week-long daily stroking sessions in the home cage. This procedure conditioned mice into a quiescent or inactive state upon gentle stroking, akin to pets being groomed (Fig. 3, A and B, and movies S1 to S3). Second, to increase the motivational drive to obtain rewards associated with pleasant touch (*27*), we performed eight conditioning sessions (one session per

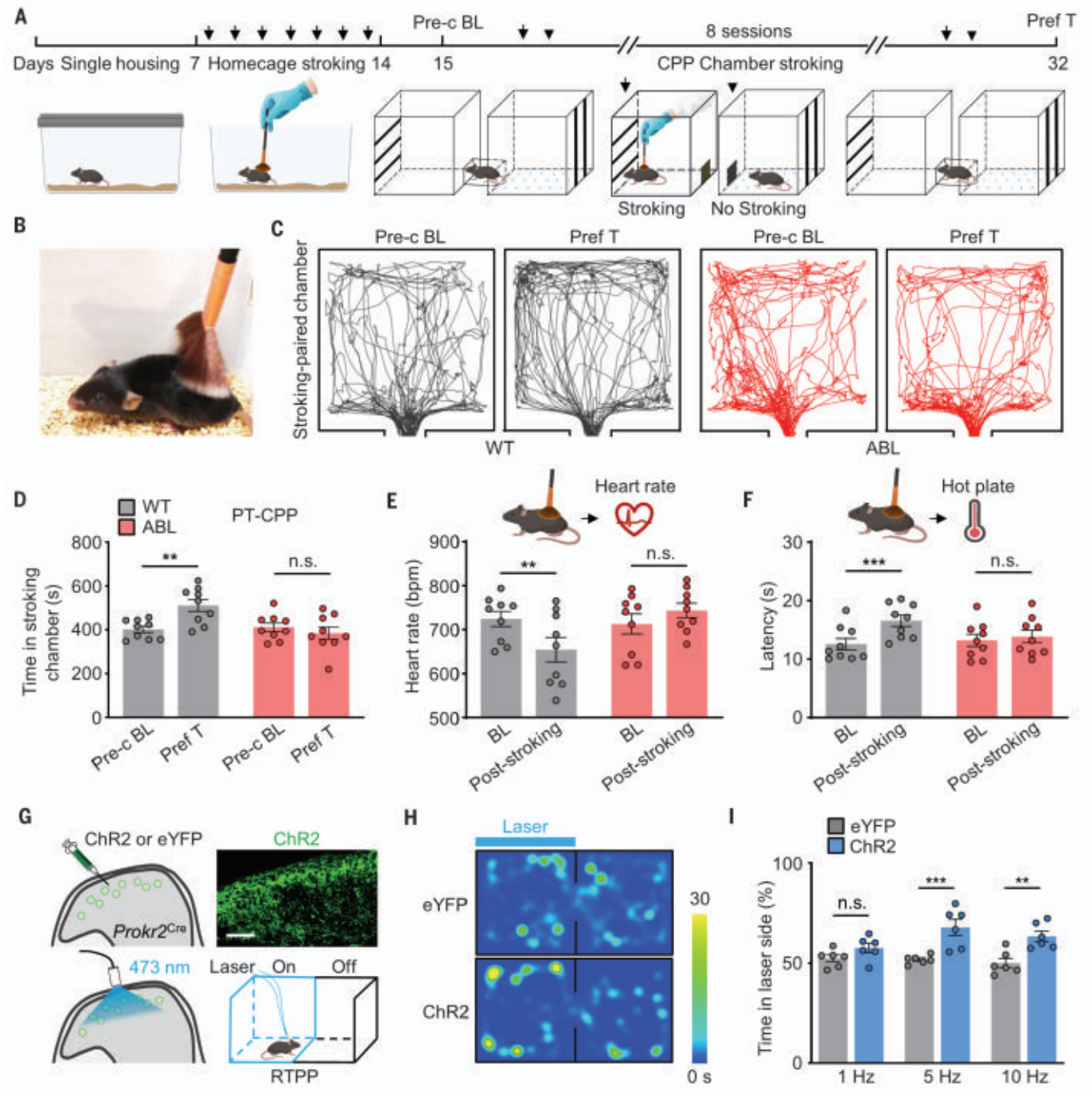


Fig. 3. PROKR2 neurons transmit pleasant touch sensation. (**A**) A schematic diagram of the experimental procedure for PT-CPP. (**B**) Photo of a mouse being stroked with a soft brush. (**C** and **D**) Representative trajectory plots (C) and time spent (D) in stroking-paired chamber for WT and ABL mice in the preconditioning baseline (Pre-c BL) and preference test (Pref T). Pre-c BL versus Pref T: $P < 0.01$ for WT and $P = 0.6822$ for ABL mice. (**E** and **F**) Heart rate (E) and thermal pain threshold (F) in the baseline and poststroking test. Heart rate, BL versus poststroking: $P < 0.01$ for WT and $P = 0.1563$ for ABL mice; hot plate, BL versus poststroking: $P < 0.001$ for WT and $P = 0.5206$ for ABL mice. bpm, beats per minute. (**G**) Schematic of intraspinal injection of Cre-dependent AAV virus expressing ChR2-eYFP or eYFP (top left) and image showing ChR2-eYFP in the cervical cord (top right). Schematic of real-time place preference test (RTPP) (bottom). Scale bar, 50 μm. (**H** and **I**) Representative heatmap (5 Hz) (H) and the percentage of time spent (I) in the chamber paired with the laser in the RTPP test. $n = 9$ [(D) to (F)]; $n = 6$ (I). Statistics by two-way repeated measures analysis of variance (ANOVA) followed by Bonferroni's multiple comparisons test [(D), (E), and (F)] or by two-way ANOVA followed by Bonferroni's multiple comparisons test (I). $^{**}P < 0.01$; $^{***}P < 0.001$; n.s., not significant. Error bars indicate SEMs.

2 days) using an unbiased two-chamber pleasant touch–conditioned place preference (PT-CPP) apparatus in which WT mice displayed no preference for either chamber (Fig. 3A, fig. S4B, and movie S4). On the test day, we performed a PT-CPP test to evaluate whether mice would spend more time in the chamber paired with gentle stroking than without. Both male and female mice developed PT-CPP as they spent significantly more time in the chamber paired with a soft brush (fig. S4, A, C, and D). These results demonstrate that gentle stroking in mice encodes the positive valence or hedonic value. In marked contrast to WT mice, ABL mice completely failed to show PT-CPP (Fig. 3, C and D), which indicates a profound loss of pleasant touch sensation. If PROKR2 neurons convey pleasant touch, their direct activation—in the absence of either primary afferent input or behavioral conditioning or context—should be positively reinforcing. We injected Cre-dependent channelrhodopsin-2 (ChR2) or enhanced YFP (eYFP) virus intraspinally into *Prokr2*^{Cre} mice followed by the real-time place preference test (RTPP) using optogenetics (Fig. 3G). Consistently, *Prokr2*$^{Cre\text{-}ChR2}$ mice preferred the chamber paired with photostimulation of PROKR2 neurons (5 and 10 Hz), whereas *Prokr2*$^{Cre\text{-}eYFP}$ mice showed no preference (Fig. 3, H and I). Together, these results demonstrate that PROKR2 neurons encode positive valence and rewarding value.

Apart from pleasant sensation (*1*), gentle stroking on the hairy skin of humans decreases heart rate (*28*) and pain sensation (*29*), reflecting a soothing state with a reduced level of stress (e.g., increasing the threshold for thermal pain). To ascertain whether mice might develop similar physiological changes, we measured the heart rate, thermal pain, and stress hormones after gentle stroking. Mice showed significantly reduced heart rates, increased thermal pain thresholds, and reduced plasma corticosterone levels (fig. S4, E to G). Consistent with behavioral studies, ABL mice displayed no significant reduction of the heart rate or analgesic effect in response to gentle stroking (Fig. 3, E and F). The lack of behavioral and physiological changes in ABL mice further demonstrates the crucial role of PROKR2 neurons in conveying pleasant touch.

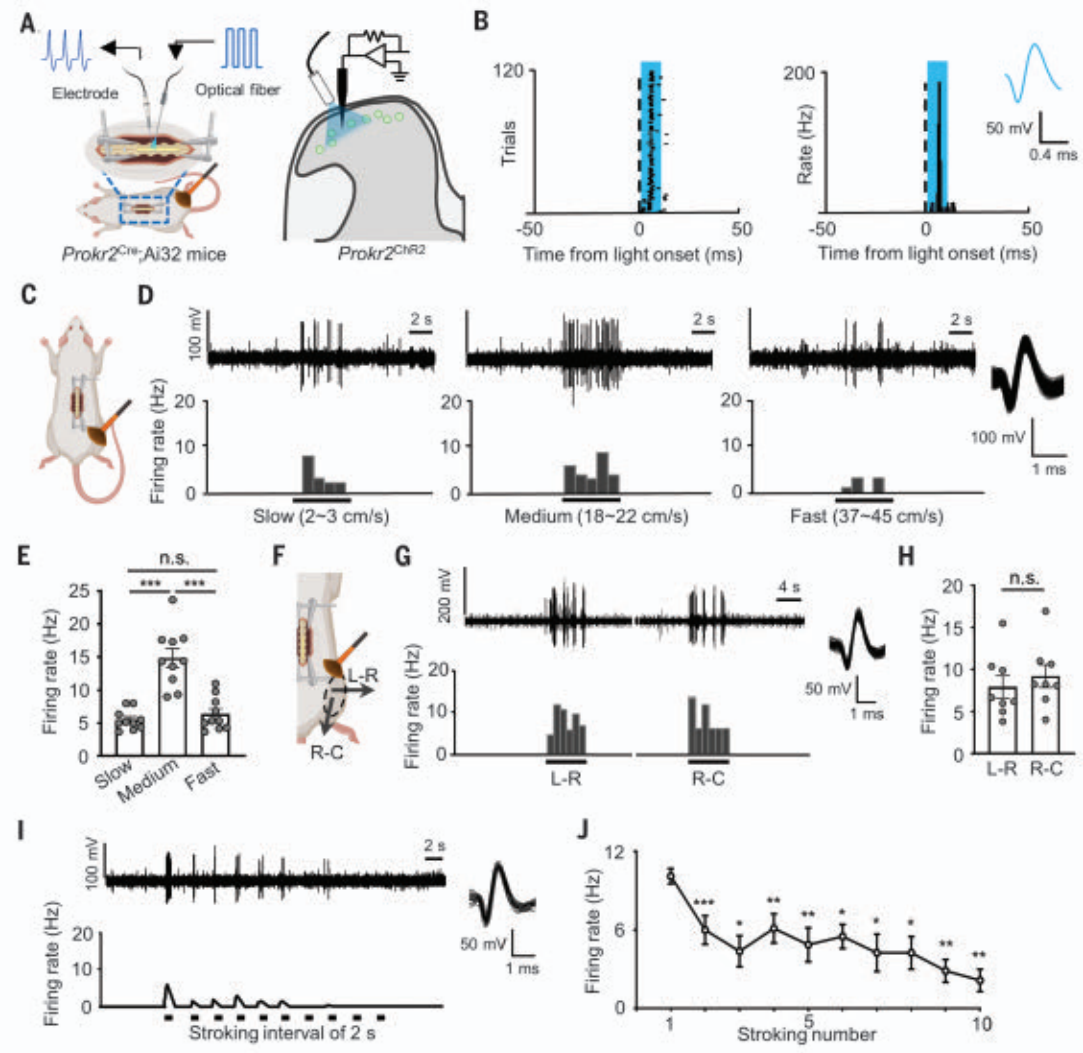


Fig. 4. PROKR2 neurons display characteristic features in response to gentle stroking. (**A**) (Left) Dorsal view of an anesthetized *Prokr2*^{Cre};Ai32 mouse showing the position of optical fiber and electrode implanted in the lumbar spinal cord. (Right) Cross-sectional view of the lumbar spinal cord showing optogenetic tagging of *Prokr2*^{ChR2} neurons in lamina II with blue lights on. (**B**) Example response of a PROKR2-expressing neuron to blue light activation. (Left) Spike raster showing multiple trials of laser stimulation at 1 Hz. (Right) The firing rate of one opto-tagged neuron within 10-ms light pulses. (Inset) Waveform on expanded time scale. Blue bar indicates a light pulse. (**C** to **E**) Schematic of brush stroking across the receptive field (C), representative traces (top) and corresponding peristimulus time histograms (PSTHs) (1-s bin) (bottom) (D), and firing rate (E) of *Prokr2*^{ChR2} neurons in response to a soft brush stroking moving at different speeds (slow, 2 to 3 cm/s; medium, 18 to 22 cm/s; fast, 37 to 45 cm/s) across the receptive field of hindlimb hairy skin. Dots represent spike rate of a single trial from individual neurons (E). (**F** to **H**) Schematic of brush stroking in different directions (F) (L-R: from left to right; R-C: from rostral to caudal), representative traces (top) and corresponding PSTHs (1-s bin) (bottom) (G), and firing rate (H) of spinal *Prokr2*^{ChR2} neurons in response to stroking in different directions. Brush stroking was applied across the receptive field of the hindlimb in different directions at a speed of 18 to 22 cm/s. (**I** and **J**) Representative trace (top) and corresponding PSTH (1-s bin) (bottom) (I) and firing rate (J) of spinal *Prokr2*^{ChR2} neurons in response to 10 repeated brush stroking stimuli with intervals of 2 s. (Inset) Superimposed waveforms on expanded time scale representing spikes of *Prokr2*^{ChR2} neurons evoked by stroking in (D), (G), and (I). n = 8 to 10 neurons from three to four mice. Statistics by one-way repeated measures ANOVA followed by Bonferroni post hoc [(E) and (J)] or by paired t test (H). $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; n.s., not significant. Error bars indicate SEMs.

Neurophysiological features of PROKR2 neurons in response to varying stimuli

To interrogate neural correlates of pleasant touch, we performed in vivo extracellular recording of the response of spinal PROKR2 neurons to gentle stroking using a soft brush moving across the receptive field in the hairy skin of the hindlimb of *Prokr2*^{Cre};Ai32 mice expressing ChR2-eYFP (Fig. 4A). Opto-tagged spinal neurons were classified as spinal *Prokr2*^{ChR2} neurons if they displayed reliable action potentials with a short response latency (10 ms) upon the delivery of a brief pulse of blue light (Fig. 4B) (*30*). Notably, the response of *Prokr2*^{ChR2} neurons was most vigorous when the hairy

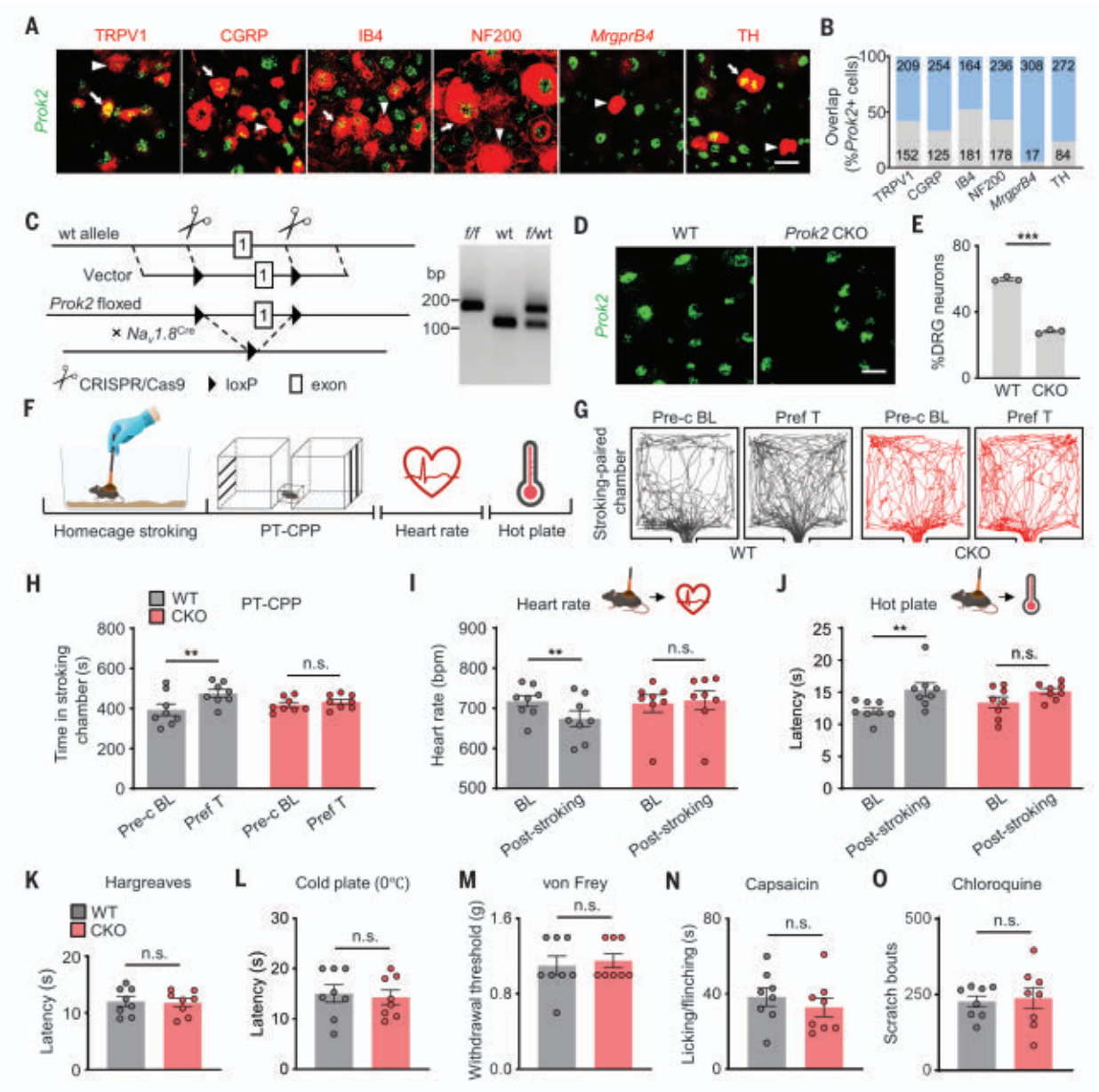


Fig. 5. Conditional deletion of *Prok2* in sensory neurons abolishes pleasant touch sensation. (**A**) Double staining of *Prok2* with various markers in DRG neurons. Arrows indicate double-stained cells, and arrowheads indicate *Prok2* only. Scale bar, 20 μm. (**B**) Overlapping percentages of (A). Gray indicates double-stained ratio, and blue indicates *Prok2* only. (**C**) (Left) Schematic of targeting strategy for generating *Prok2* floxed (*Prok2*$^{f/f}$) mice, which were mated with *Na$_v$1.8*^{Cre} mice to generate *Prok2*$^{f/f}$;*Na$_v$1.8*^{Cre} mice or *Prok2* CKO mice. (Right) Gel electrophoresis of genotyping polymerase chain reaction (PCR) from *Prok2*$^{f/f}$, wt/wt, and *Prok2*$^{f/wt}$ samples. (**D**) Expression of *Prok2* in DRGs of the control (*Prok2*$^{f/f}$, WT) and *Prok2* CKO mice (*Prok2*$^{f/f}$;*Na$_v$1.8*^{Cre}, CKO). Scale bar, 20 μm. (**E**) Quantification of (D). (**F**) A schematic of the experimental procedure. Parallel symbols indicate that PT-CPP, heart rate, and hot plate tests were performed independently after single-housing and home-cage stroking. (**G** and **H**) Representative trajectory plots (G) and time spent (H) in stroking-paired chamber for WT and CKO mice in the preconditioning baseline (Pre-c BL) and preference test (Pref T). Pre-c BL versus Pref T: $P < 0.01$ for WT and $P = 0.9552$ for CKO mice. (**I** and **J**) Heart rate test (I) and hot plate test (J) in WT and CKO mice. Heart rate, BL versus poststroking: $P < 0.01$ for WT and $P = 0.8737$ for CKO mice; hot plate, BL versus poststroking: $P < 0.01$ for WT and $P = 0.1531$ for CKO mice. (**K** to **O**) Comparable latencies in Hargreaves test (K) and cold plate test (L), withdraw threshold in von Frey test (M), licking or flinching time induced by capsaicin (2 g, i.pl.) (N), and scratching numbers induced by chloroquine injection (200 μg, i.d.) (O) between WT and CKO mice. $n = 3$ (E); $n = 8$ [(H) to (O)]. Statistics by unpaired t test [(E) and (K) to (O)] or by two-way repeated measures ANOVA followed by Bonferroni's multiple comparisons test [(H), (I), and (J)]. $^{**}P < 0.01$; $^{***}P < 0.001$; n.s., not significant. Error bars indicate SEMs.

skin of mice was brushed at a slowly moving speed (18 to 22 cm/s); however, their firing became sluggish if the stroking speed was slower (2 to 3 cm/s) or faster (37 to 45 cm/s) (Fig. 4, C and D). The mean firing rates at three brushing speeds exhibited an inverted U shape (Fig. 4E), reminiscent of that of human CT fibers (*17*, *18*). This prompted us to examine whether other features of *Prokr2*^{ChR2} neurons may resemble the hallmarks of human CT fibers (*1*, *16*, *17*, *19*). One signature of CT fibers is that they show no preference for orientation in the receptive field (*19*). *Prokr2*^{ChR2} neurons displayed comparable firing rates, irrespective of stroking directions (e.g., from rostral to caudal or from left to right; Fig. 4, F to H). Another distinctive feature is fatigue to the repetition of brush stroking, referring to gradually attenuated responses to repeated stroking stimuli within seconds (*1*, *16*, *17*, *19*). This feature distinguishes unmyelinated CT fibers from myelinated low-threshold mechanoreceptors (*18*) and has also been observed in cutaneous C mechanoreceptor afferents in rats (*31*) and cats (*32*). The firing rate of *Prokr2*^{ChR2} neurons also exhibited fatigue to repeated brush stroking and was reduced by 40.7 to 78.8% within 10 series of successive brush stroking (2-s interval) (Fig. 4, I and J). Consistent with a recent study showing the presence of CT afferents in human glabrous skin (*33*), PROKR2 neurons also responded to gentle stroking applied to the glabrous skin (fig. S5A), indicating the presence of PROK2 fibers in the glabrous skin.

At last, we examined several nongentle stroking–related stimuli that could elicit the response of CT fibers (*16*, *34*). Consistently, *Prokr2*^{ChR2} neurons responded to punctate stimulation (von Frey filament at 0.07 g or 0.7 mN), pinprick stimulation, and cooling temperature (30° to 15°C) (fig. S5, B to D).

Coding of pleasant touch by PROK2 in sensory neurons

We next examined *Prok2* expression in several types of dorsal root ganglion (DRG) neurons using RNAscope combined with IHC. *Prok2* partially overlaps to varying degrees with TRPV1 and calcitonin gene–related peptide (CGRP), which are expressed in nociceptive neurons; IB4; NF200, a neurofilament expressed predominantly in large-diameter neurons; *MrgprB4*; and tyrosine hydroxylase (TH) expressed in some C–low-threshold mechanoreceptors (C-LTMRs) (*35*) (Fig. 5, A and B, and fig. S6F). Overall, ~59.6% of DRG neurons express *Prok2*. To ascertain whether *Prok2*-expressing primary afferents form monosynaptic contacts with PROKR2 neurons, we performed rabies virus–mediated retrograde tracing in the spinal cord of *Prokr2*^{Cre} mice (*25*) (fig. S6, A and B). Examination of *Prok2* expression with retrograde transported glycoprotein (G)–deleted rabies virus (RVdG) that labels the input neurons with dsRed in DRGs revealed that ~70% of the input neurons were colabeled with *Prok2* (fig. S6, C to E). Only a small portion of CGRP fibers and minimal TRPV1 or TH fibers formed direct contacts with PROKR2 neurons, whereas no *MrgprB4* and IB4 afferents were found to form monosynaptic contacts with PROKR2 neurons (fig. S6, C and D). However, it should be noted that a lack of inputs from TH or IB4 afferents could be because of their resistance to rabies virus infection (*36*). Further, we show that the sizes of *Prok2* input neurons normally are in the 200-to-400-μm^2 range, which indicates that they are small DRG neurons (fig. S6E).

To determine the role of PROK2 in pleasant touch, we generated mice harboring a floxed allele of *Prok2* using the gene targeting strategy (Fig. 5C). Floxed *Prok2* mice were bred with *Na$_v$1.8*^{Cre} mice, which express the Cre recombinase in small nociceptive sensory neurons, to delete *Prok2* in *Na$_v$1.8*^{Cre} neurons of DRGs (Fig. 5C). The expression of *Prok2* was reduced by 52.8% without affecting IB4, CGRP, or *MrgprB4* in DRGs and IB4 or CGRP central fibers in the lumbar cord (Fig. 5, C to E, and fig. S7, A to D). Mice with conditional knockout of *Prok2* in DRGs, referred to as *Prok2* CKO mice, failed to show PT-CPP (Fig. 5, F to H). By contrast, they showed normal pain and itch behaviors (Fig. 5, K to O). Furthermore, *Prok2* CKO mice did not show significant changes in the heart rate and thermal pain thresholds after gentle stroking (Fig. 5, I and J). CT and C-LTMRs are insensitive to capsaicin, which activates TRPV1 fibers (*37*, *38*). This prompted us to explore the role of TRPV1 fibers in pleasant touch. Ablation of the central terminals of TRPV1 fibers with intrathecal resiniferatoxin (RTX), a potent TRPV1 agonist, abolished neurogenic pain elicited by capsaicin and markedly attenuated mechanical itch that is dependent on Aβ-LTMR and

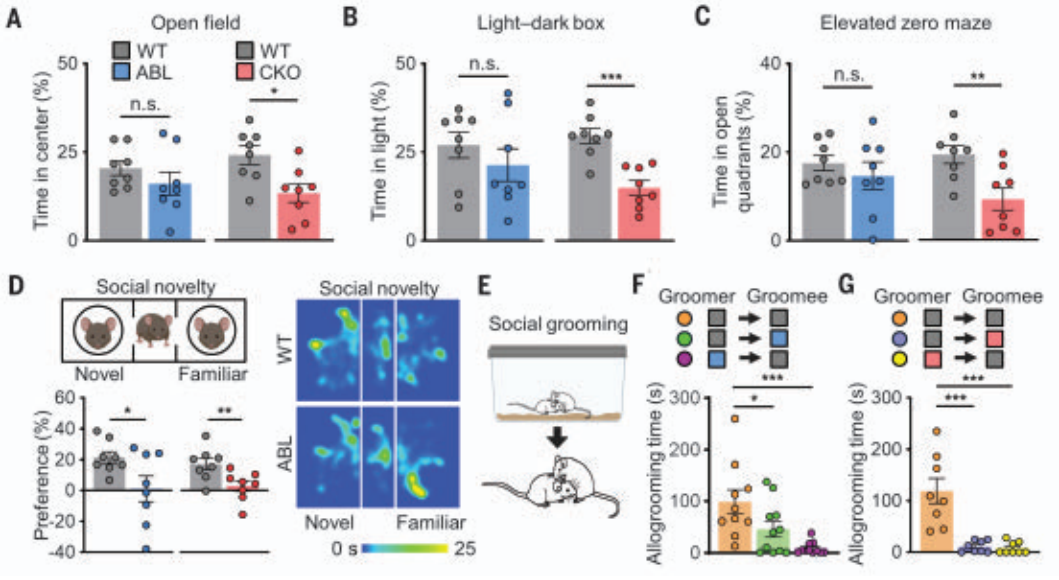


Fig. 6. Profound impairments of PROR2-PROKR2 mutant mice in stress response and prosocial behaviors. (**A** to **C**) Stress and anxiety-like behavioral tests. Shown are the percentages of time spent in the center zone of the open field apparatus (A), time spent in the light-illuminated chamber of the light-dark box (B), and time spent in the open quadrants of the elevated zero maze (C). WT versus ABL: $P = 0.2795$ (A), $P = 0.3514$ (B), and $P = 0.4312$ (C); WT versus CKO: $P < 0.05$ (A), $P < 0.001$ (B), and $P < 0.01$ (C). (**D**) The three-chamber social novelty test. The preference index for the percentage of time spent exploring the chamber with an unfamiliar mouse versus a familiar mouse (left). (Right) Representative heatmaps of locomotor activity in the chambers. WT versus ABL: $P < 0.05$; WT versus CKO: $P < 0.01$. (**E** to **G**) The home-cage social grooming test. Cartoons show mouse allogrooming in the home cage (E). (F and G) Allogrooming time for each pair. Groomer-groomee (F), WT-WT pair versus WT-ABL pair: $^*P < 0.05$; WT-WT pair versus ABL-WT pair: $^{***}P < 0.001$. Groomer-groomee (G), WT-WT pair versus WT-CKO pair: $^{***}P < 0.001$; WT-WT pair versus CKO-WT pair: $^{***}P < 0.001$. $n = 8$ [(A) to (D) and (G)]; $n = 10$ to 11 (F). Statistics by unpaired t test [(A) to (D)] or one-way ANOVA followed by Bonferroni post hoc [(F) and (G)]. $^*P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; n.s., not significant. Error bars indicate SEMs.

C fiber inputs and GRP and GRPR neurons (*22*, *25*) (fig. S8, A to D). Notably, RTX treatment had no effect on PT-CPP (fig. S8, E to G).

Profound deficits in stress response and prosocial behaviors of PROK2 mutant mice

Pleasant social touch affords tremendous emotional and psychological benefits by indirectly activating the endogenous reward-pleasure circuits to release a plethora of neuropeptides and neurotransmitters that encode antistress, positive hedonic, and prosocial value (*1*, *7*, *39*–*42*). We sought to examine whether a loss of pleasant touch developmentally or in adult mice may result in abnormal stress responses. *Prok2* CKO mice spent significantly less time than their WT littermates in the center zone of the open-field apparatus, in the light-illuminated chamber of the light-dark box, and in the open quadrants of the elevated zero maze apparatus (Fig. 6, A to C). By contrast, the time that ABL mice spent in these areas did not differ from that of their WT littermates (Fig. 6, A to C). We further assessed social novelty recognition of mutant mice using the three-chamber social interaction test. Unlike WT mice, neither *Prok2* CKO nor ABL mice displayed a preference for a newly introduced mouse, demonstrating severe deficits in social novelty recognition (Fig. 6D and fig. S9).

Finally, we evaluated whether *Prok2* CKO or ABL mice might have deficits in social touch behavior. To examine this, we first monitored spontaneous social interactions between C57BL-6J paired adult mice in the home cage, which subserves a more naturalistic environment, and observed frequent social or allogrooming behaviors unrelated to sexual or aggressive and/or conflict activity (Fig. 6E and movie S5). In sharp contrast to the paired WT-WT conspecifics, social grooming of WT mice toward ABL mice was significantly diminished (Fig. 6F). Unexpectedly, ABL mice rarely groomed WT conspecifics capable of sensing pleasant touch (Fig. 6F). These deficits were even more pronounced in the WT–*Prok2* CKO pairs (Fig. 6G).

Discussion

Using an interdisciplinary approach coupled with autonomic, neuroendocrine, electrophysiological, and behavioral criteria, our study demonstrates the crucial function of the PROK2-PROKR2 signaling pathway in pleasant touch. The profound loss of pleasant touch sensation in *Prok2* CKO mice underscores the pivotal role of PROK2 in the coding and transmitting of pleasant touch information. Our work reveals two parallel peptidergic pathways, distinguished by their capsaicin sensitivity, from the skin to the spinal cord: One conveys positive valence to PROKR2 interneurons, which constitutes an obligate circuit for sending pleasant touch to the brain via *Gpr83* neurons, and the other conveys negative valence to the brain through spinal pain- and itch-specific local microcircuits and NK1R projection neurons (*20*, *24*, *43*). We argue that PROK2 in Aβ and Aδ fibers is unlikely to be involved in pleasant touch because PROKR2 neurons receive no direct inputs from these fibers. Moreover, unlike unmyelinated C fibers that encode the specificity of sensory modalities through slow-acting neuropeptides, myelinated Aβ and Aδ fibers typically use fast-acting glutamate as a neurotransmitter to relay information (*22*). We establish the function of PROK2 in encoding affective properties of gentle stroking; however, understanding how Aβ and Aδ LTMR fibers contribute to the transmission of discriminative properties of gentle touch and stroking will require further studies. Although PROK2 fibers that synapse with PROKR2 neurons could be the equivalent of human CT fibers, it is not feasible to examine their conduction velocities or cutaneous innervation patterns in isolation. Nonetheless, notable neurophysiological features shared by PROKR2 neurons and human CT fibers indicate that a subset of PROK2 fibers is equivalent to CT fibers. The response of PROKR2 neurons to cooling raises the possibility that PROKR2 neurons may be a convergent node for integrating different kinds of cutaneous information that encodes positive hedonic valence (e.g., pleasantness of cooling), akin to GRPR neurons subserving as a convergent node for mechanical and chemical itch (*25*). These data further support the neuropeptide code model that somatosensory modalities with slow response kinetics are encoded by neuropeptides in sensory neurons and conveyed by respective spinal microcircuits that can be defined and identifiable through specific GPCR expression (*22*). The development of the PT-CPP test that entails inference of pleasant touch in an unbiased manner overcomes a major obstacle in the interrogation of molecular underpinnings and neural circuits of pleasant touch. Together with the ethologically relevant seminaturalistic social grooming paradigm, we provide an avenue to unravel mechanisms by which the need for affective touch drives social attachment and affiliative behaviors.

This study has important clinical implications. The heightened stress and anxiety-like behaviors of *Prok2* CKO but not ABL mice are attributable to a lack of pleasant touch during a critical period in development, which supports prior studies showing that early tactile experience is more instrumental in shaping the resilience of offspring against stressful events (*44*, *45*). Accordingly, ABL mice are less vulnerable to stress. In light of the challenge in the study of the developmental role of pleasant touch because of the multisensory nature of parental care (*46*), *Prok2* CKO mice might serve as an invaluable animal model for assessment of the long-term effects of deprivation of maternal or caregiving nurturing touch on offsprings' development and health (*47*). A dearth of pleasant touch sensation could dysregulate prosocial neuropeptide expression in the brain and thereby impede social recognition and interactions. The failure of mutant mice to recognize unfamiliar conspecifics indicates a crucial role of pleasant touch in social recognition and social memory known to be important for social bonding (*48*). Additionally, the marked avoidance of the social touch phenotype is reminiscent of some early traits of ASD (*10*, *14*, *49*). Further analysis of prosocial behavioral impairment in *Prok2* CKO mice may offer additional insights into the etiology of certain neurodevelopmental and affective disorders that hinder social interactions and affiliative behaviors. The inability of mutant mice to groom WT conspecifics stresses the importance of PROK2 signaling in mediating synchronous and bidirectional communication of mutually beneficial social information through reciprocal tactile contacts (*6*). Finally, the observation that the firing features of PROKR2 neurons recapitulate the hallmarks of human CT fibers reinforces the notion that neural mechanisms of pleasant touch are conserved between humans and rodents. Conceivably, a deficiency of PROK2-PROKR2 signaling might result in social and emotional impairments that could lead to social isolation, anxiety, and mental disorders.

REFERENCES AND NOTES

1. F. McGlone, J. Wessberg, H. Olausson, *Neuron* **82**, 737–755 (2014).
2. V. E. Abraira, D. D. Ginty, *Neuron* **79**, 618–639 (2013).
3. I. Morrison, L. S. Löken, H. Olausson, *Exp. Brain Res.* **204**, 305–314 (2010).
4. A. Montagu, *Touching: The Human Significance of the Skin* (Harper Collins, ed. 3, 1986).
5. J. Panksepp, *Affective Neuroscience: The Foundations of Human and Animal Emotions* (Affective Science Series, Oxford Univ. Press, 1998).
6. M. J. Hertenstein, J. M. Verkamp, A. M. Kerestes, R. M. Holmes, *Genet. Soc. Gen. Psychol. Monogr.* **132**, 5–94 (2006).
7. R. I. M. Dunbar, *Neurosci. Biobehav. Rev.* **34**, 260–268 (2010).
8. F. B. M. de Waal, M. Suchak, *Phil. Trans. R. Soc. B* **365**, 2711–2722 (2010).
9. E. O. Wilson, *Sociobiology: The New Synthesis* (Harvard Univ. Press, 1975).
10. C. J. Cascio, D. Moore, F. McGlone, *Dev. Cogn. Neurosci.* **35**, 5–11 (2019).
11. H. F. Harlow, *Am. Psychol.* **13**, 673–685 (1958).
12. R. A. Spitz, *Psychoanal. Study Child* **1**, 53–74 (1945).
13. J. Bowlby, *Bull. World Health Organ.* **3**, 355–533 (1951).
14. A. C. Voos, K. A. Pelphrey, M. D. Kaiser, *Soc. Cogn. Affect. Neurosci.* **8**, 378–386 (2013).
15. M. D. Thye, H. M. Bednarz, A. J. Herringshaw, E. B. Sartin, R. K. Kana, *Dev. Cogn. Neurosci.* **29**, 151–167 (2018).
16. Å. Vallbo, L. Löken, J. Wessberg, in *Affective Touch and the Neurophysiology of CT Afferents*, H. Olausson, J. Wessberg, I. Morrison, F. McGlone, Eds. (Springer, 2016), pp. 1–30.
17. L. S. Löken, J. Wessberg, I. Morrison, F. McGlone, H. Olausson, *Nat. Neurosci.* **12**, 547–548 (2009).
18. A. B. Vallbo, H. Olausson, J. Wessberg, *J. Neurophysiol.* **81**, 2753–2763 (1999).
19. J. Wessberg, H. Olausson, K. W. Fernström, A. B. Vallbo, *J. Neurophysiol.* **89**, 1567–1575 (2003).
20. S. Choi *et al.*, *Nature* **587**, 258–263 (2020).
21. S. Vrontou, A. M. Wong, K. K. Rau, H. R. Koerber, D. J. Anderson, *Nature* **493**, 669–673 (2013).
22. Z. F. Chen, *Nat. Rev. Neurosci.* **22**, 758–776 (2021).
23. Y. G. Sun, Z. F. Chen, *Nature* **448**, 700–703 (2007).

24. Y. G. Sun *et al*., *Science* **325**, 1531–1534 (2009).
25. S. Chen *et al*., *Nat. Commun.* **11**, 5074 (2020).
26. E. B. Keverne, N. D. Martensz, B. Tuite, *Psychoneuroendocrinology* **14**, 155–161 (1989).
27. C. L. Cunningham, C. M. Gremel, P. A. Groblewski, *Nat. Protoc.* **1**, 1662–1670 (2006).
28. R. Pawling, P. R. Cannon, F. P. McGlone, S. C. Walker, *PLOS ONE* **12**, e0173457 (2017).
29. L. L. Meijer, C. Ruis, M. J. van der Smagt, E. J. A. Scherder, H. C. Dijkerman, *J. Neuropsychol.* **16**, 38–53 (2022).
30. J. P. Fadok *et al*., *Nature* **542**, 96–100 (2017).
31. J. W. Leem, W. D. Willis, J. M. Chung, *J. Neurophysiol.* **69**, 1684–1699 (1993).
32. A. Iggo, *J. Physiol.* **152**, 337–353 (1960).
33. R. H. Watkins *et al*., *J. Neurophysiol.* **125**, 232–237 (2021).
34. M. Nordin, *J. Physiol.* **426**, 229–240 (1990).
35. L. Li *et al*., *Cell* **147**, 1615–1627 (2011).
36. G. W. Albisetti *et al*., *J. Neurosci.* **37**, 10358–10371 (2017).
37. F. McGlone, S. Walker, R. Ackerley, in *Affective Touch and the Neurophysiology of CT Afferents*, H. Olausson, J. Wessberg, I. Morrison, F. McGlone, Eds. (Springer, 2016), pp. 265–282.
38. M. J. Caterina *et al*., *Nature* **389**, 816–824 (1997).
39. K. Uvnäs-Moberg, L. Handlin, M. Petersson, *Front. Psychol.* **5**, 1529 (2015).
40. J. Panksepp, B. H. Herman, T. Vilberg, P. Bishop, F. G. DeEskinazi, *Neurosci. Biobehav. Rev.* **4**, 473–487 (1980).
41. H. Yu *et al*., *Neuron* **110**, 1051–1067.e7 (2022).
42. Y. Tang *et al*., *Nat. Neurosci.* **23**, 1125–1137 (2020).
43. D. Mu *et al*., *Science* **357**, 695–699 (2017).
44. M. J. Meaney, *Annu. Rev. Neurosci.* **24**, 1161–1192 (2001).
45. A. Kaffman, M. J. Meaney, *J. Child Psychol. Psychiatry* **48**, 224–244 (2007).
46. C. Dulac, L. A. O'Connell, Z. Wu, *Science* **345**, 765–770 (2014).
47. F. S. Hall, *Crit. Rev. Neurobiol.* **12**, 129–162 (1998).
48. H. Walum, L. J. Young, *Nat. Rev. Neurosci.* **19**, 643–654 (2018).
49. K. L. Bales *et al*., *Neurosci. Biobehav. Rev.* **95**, 202–219 (2018).

ACKNOWLEDGMENTS

We thank X. Liu for technical help, Y. Zhang and F. Gao for help with viral tracing, Q. Ma and M. Goulding for providing *Lbx1*^{Flpo} and *Tau*$^{ds\text{-}DTR}$ lines, L. Setton and T. McGrath for help with the initial design of the stroking test, and R. Bardoni for discussion. **Funding:** The project has been supported by NIH grants 1R01AR056318-06 and R01NS094344 (to Z.-F.C.). **Author contributions:** B.L. performed immunostaining, genetic ablation, and electrophysiological studies. L.Q. and K.L. developed the PT-CPP test and performed in vivo extracellular recording. J.L. performed RNAscope and double staining. T.J.P.-A. participated in the behavioral tests. B.L., L.Q., and Z.-F.C. wrote the manuscript. Z.-F.C. conceived and supervised the project. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abn2479
Materials and Methods
Figs. S1 to S9
References (*50*–*69*)
MDAR Reproducibility Checklist
Movies S1 to S5

13 November 2021; accepted 24 March 2022
10.1126/science.abn2479

ZEOLITES

In situ imaging of the sorption-induced subcell topological flexibility of a rigid zeolite framework

Hao Xiong[1]†, Zhiqiang Liu[2]†, Xiao Chen[1]*, Huiqiu Wang[1], Weizhong Qian[1], Chenxi Zhang[1]*, Anmin Zheng[2], Fei Wei[1]*

The crystallographic pore sizes of zeolites are substantially smaller than those inferred from catalytic transformation and molecular sieving capabilities, which reflects flexible variation in zeolite opening pores. Using in situ electron microscopy, we imaged the straight channels of ZSM-5 zeolite with benzene as a probe molecule and observed subcell flexibility of the framework. The opening pores stretched along the longest direction of confined benzene molecules with a maximum aspect change of 15%, and the *Pnma* space group symmetry of the MFI framework caused adjacent channels to deform. This compensation maintained the stability and rigidity of the overall unit cell within 0.5% deformation. The subcell flexibility originates mainly from the topologically soft silicon-oxygen-silicon hinges between rigid tetrahedral SiO_4 units, with inner angles varying from 135° to 153°, as confirmed by ab initio molecular dynamics simulations.

Zeolite pores, which have diameters ranging from ~0.3 to ~1.3 nm, dictate their molecular sieving properties and control access to internal sites that are catalytically active or preferred for binding by sorbates (*1*–*4*). By adjusting the size and shape of pore openings, the size and shape of adsorbing molecules that fit inside the pores can be selected, thereby excluding larger molecules. This effect is exploited in selective chemical conversion (*4*–*9*). However, the effective pore sizes calculated from the crystallographic structures are substantially smaller than those inferred from catalytic transformation and molecular sieving capabilities (*8*–*10*). It has long been speculated that this discrepancy reflects in part a flexible deformation of zeolite pores.

Nonetheless, the flexible deformation of zeolite pores is rarely reported. Unlike flexible metal-organic frameworks (MOFs) with long and soft organic linkers (*11*–*13*), zeolite materials (with an elastic modulus in the range of 50 to 100 GPa) macroscopically behave as rigid and fragile materials in most applications (*14*–*16*). Findings related to zeolite flexibility have included minor variations in cell parameters (<10 pm) and a symmetry transformation from monoclinic to orthorhombic as a function of temperature and sorbates (*17*–*23*).

Direct measurement of pore sizes is, in principle, possible by diffraction, such as x-ray diffraction and electron diffraction. Nonetheless, the derived Debye-Waller factors for zeolites usually compensate for the complexity of zeolite crystal structures and various other model deficiencies, including inhomogeneity effects. Furthermore, averaged cell-scale results cannot indicate the local structure of deformed opening pores caused by successive twisted tetrahedrons. Crystal dynamics simulation techniques have been applied to trace the motion and vibration of specific atoms within zeolite frameworks, and some progress toward describing framework flexibility has been made, but supporting experimental evidence has been lacking (*2*, *9*, *24*–*26*). Thus, new experimental technology with ultrahigh spatial resolution is needed to reveal the subcell local structural evolution of the zeolite framework when guest molecules break through the pore size limit, especially in practical applications.

Integrated differential phase contrast scanning transmission electron microscopy (iDPC-STEM) can reveal the local structures of zeolites and image the confined molecules inside them in real space at an atomic scale (*27*–*30*) because of its excellent capability for imaging light and heavy elements together (*31*, *32*). Here, we combined iDPC-STEM imaging with an in situ atmosphere system to monitor the molecular phase transition and corresponding geometric variations of opening pores during the benzene adsorption-desorption process in real time. We studied the straight channel (5.3 Å × 5.6 Å) of ZSM-5 (MFI-type) zeolite as the imaging window and used benzene, which has a kinetic diameter of 5.85 Å, as a probe molecule. We observed the phase transition of confined benzene molecules and resolved the varying atomic structure of the MFI framework. Thus, we successfully observed the subcell flexibility of the zeolite framework, investigated the local deformation of zeolite channels, and monitored the dynamic evolution process when guest molecules enter or exit the zeolite framework. The results will enable us to better understand the topologically flexible structural characteristics of zeolites and the intrinsic mechanism of molecular diffusion in microporous materials.

[1]Beijing Key Laboratory of Green Chemical Reaction Engineering and Technology, Department of Chemical Engineering, Tsinghua University, Beijing 100084, China. [2]State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, National Center for Magnetic Resonance in Wuhan, Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences, Wuhan 430071, China.
***Corresponding author. Email: chenx123@mail.tsinghua.edu.cn (X.C.); cxzhang@mail.tsinghua.edu.cn (C.Z.); wf-dce@mail.tsinghua.edu.cn (F.W.)**
†These authors contributed equally to this work.

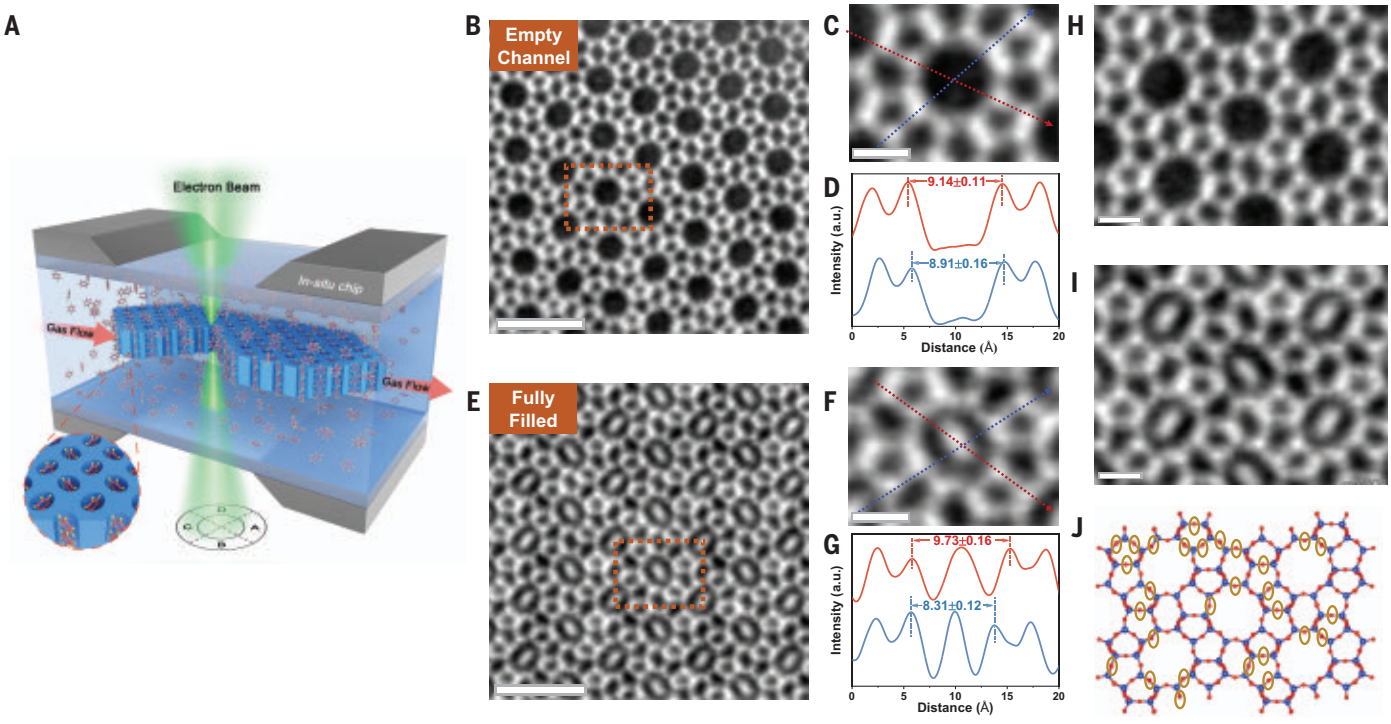


Fig. 1. Imaging the huge local deformation of zeolite channels in the in situ STEM system during benzene adsorption. (**A**) Schematic of the in situ STEM experimental setup. The benzene adsorption-desorption process was carried out in an enclosed environment chamber made of Si_3N_4 solid membranes in an atmosphere system. (**B** to **D**) iDPC-STEM images and profile analysis of an empty MFI zeolite specimen from the [010] projection. The magnified image shows a nearly round pore. The profile analysis gives the aspect ratio of D_{max} to D_{min}, which was used to estimate the local deformation of straight channels. (**E** to **G**) iDPC-STEM images and profile analysis of a saturated benzene@MFI specimen from the [010] projection at 473 K and 450 torr (90% N_2 and 10% benzene). The 10-ring opening pore shows a distinctly elliptical shape stretched along the parallel direction of the benzene plane. (**H** to **J**) Further comparison of imaging differences between the empty (H) and saturated states (I) of benzene@MFI, showing clearer imaging of O atoms in the MFI framework, which are highlighted in (J). Forty opening pores were measured to obtain each statistical result. Scale bars, 2 nm [(B) and (E)], 500 pm [(C), (F), (H), and (I)].

In situ observation of subcell deformation in zeolite channels

ZSM-5 zeolites are composed of corner-sharing tetrahedral TO_4 (T: Si, Al) and have a cross-linked channel system formed of straight channels (5.3 Å × 5.6 Å) and sinusoidal channels (5.1 Å × 5.5 Å) (fig. S1). The straight channels of MFI zeolite can be directly imaged from the [010] projection at the atomic scale and are the best imaging windows for observing opening pores and confined molecules (fig. S2). Inspired by the synthesis strategy for nanosized lamellar architectures (*8*, *33*–*37*), we synthesized coffin-shaped MFI single crystals that had a smooth surface and uniform thickness (~40 nm) along the *b* axis (figs. S3 and S4). By eliminating the influence of height variance on STEM imaging, we could observe in situ both adsorbed benzene molecules and the zeolite framework in a large field of view (25.74 nm × 25.74 nm).

Figure 1A exhibits the schematic of the in situ STEM experimental setup. The experiments were performed in a spherical aberration (Cs)–corrected transmission electron microscope equipped with an enclosed environment chamber where environmental temperature, inlet gas composition, and volume flow could be regulated precisely. Full experimental details are given in the supplementary materials. Before benzene adsorption, the fresh MFI zeolite specimen was degassed under vacuum at 923 K for 5 hours. From the [010] projection, the projected positions of Si atoms and empty straight channels were imaged (Fig. 1B); the O atoms were covered by nearby Si atoms because of thermal diffuse scattering and lattice vibrations.

To better describe the varying geometries of straight channels, we defined the distances of Si_1 to Si_6 and Si_4 to Si_9 as the long axis (D_{max}, red line) and short axis (D_{min}, blue line), respectively, as the periodic parameters for straight channels (Fig. 1, C and F, and fig. S5), which is slightly different from the definition of the pore size (the distance of O atoms) of straight channels (*38*). The magnified image of empty straight channels showed a nearly round pore shape (D_{max}, 9.14 ± 0.11 Å; D_{min}, 8.91 ± 0.16 Å) with an aspect ratio of 1.02 (Fig. 1D and fig. S6), which is highly consistent with the standard structural model (fig. S5). This symmetrical and uniform shape of opening pores originates from high crystallographic symmetry and isostatic perfectly regular tetrahedra, indicating no internal stress accumulation.

After benzene adsorption at 473 K and 450 torr (10% benzene and 90% N_2), the close size match of one C_6-ring plane and the channels of ZSM-5 zeolite caused the benzene molecules to be stably confined in straight and sinusoidal channels through van der Waals interactions (*39*, *40*). Following the principle of energy optimization, the degrees of freedom of adsorbed molecules would be reduced to several states (*28*, *29*). Thus, the confined benzene molecules were "frozen" in zeolite channels with a fixed orientation and critical dimensions (7.367 Å × 6.702 Å × 3.40 Å) (*41*, *42*) that were different from the kinetic diameter in the gas phase when assumed to be a sphere (*1*).

The total overlapping projection of these benzene molecules formed an elliptical spindle with the smallest cross-sectional area of benzene in each straight channel (Fig. 1E). The unified orientation and superimposed contrast along the [010] projection enabled us to simultaneously realize the imaging of small molecules and the zeolite framework. Moreover, we observed substantial local deformation of the straight channels, from an approximate round shape to a clear elliptical shape (Fig. 1, F and G). The straight channels were stretched along D_{max} (9.73 ± 0.16 Å), which was parallel to

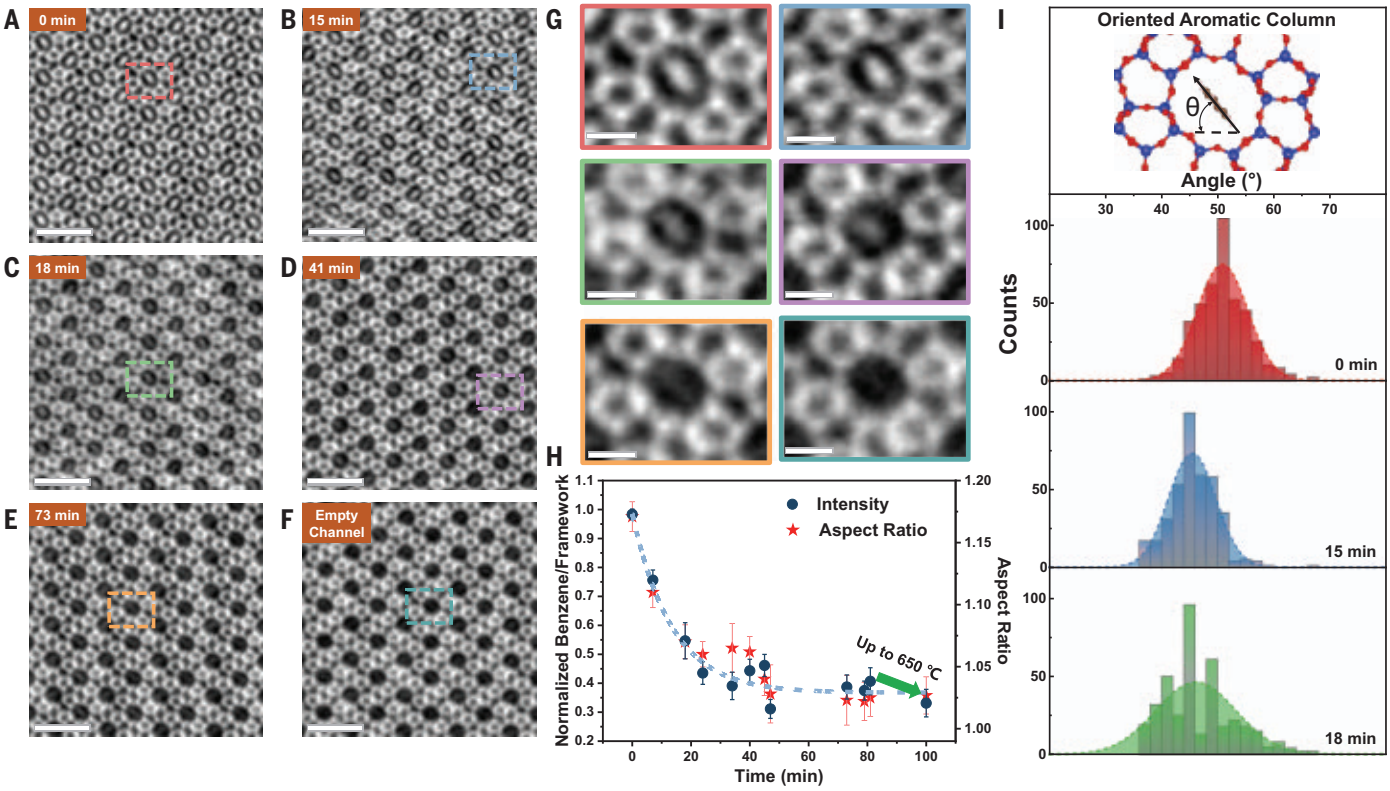


Fig. 2. Dynamic evolution of zeolite channels and corresponding host-guest interactions in benzene desorption. (**A** to **F**) iDPC-STEM snapshots of a benzene@MFI specimen along the [010] projection at different stages during the cyclic benzene desorption process. Benzene was taken up at 473 K and 450 torr (10% benzene and 90% N_2) and released under vacuum; the saturated state is denoted as 0 min. After ~80 min, the temperature was increased to 923 K to release the remaining benzene molecules. (**G**) Magnified iDPC-STEM images of straight channels extracted from (A) to (F) (marked by the corresponding colored boxes), indicating the changes in benzene contrast and channel geometry at different stages. (**H**) Evolution of the normalized benzene/zeolite framework contrast ratio and corresponding aspect ratios of Si_{10} opening pores during the in situ benzene uptake and release process. The pink stars indicate points where the aspect ratios of Si_{10} rings are correlated with the imaging contrast of confined benzene molecules, which reached 1.15 in the saturated state. Fifty opening pores were measured to obtain the error bars at each point. (**I**) Statistical distribution of benzene orientations at different stages of benzene desorption. Scale bars, 2 nm [(A) to (F)], 500 pm (G).

the C_6 plane of benzene molecules, and compressed along D_{min} (8.31 ± 0.12 Å), with an aspect ratio of 1.176 (Fig. 1F). The difference of the D_{max} and D_{min} reached 15%. The maximum pore diameter marked by Si atoms was increased by 0.6 Å from the [010] projection to match the critical dimension (6.702 Å) of benzene molecules, determined from the smallest cross section. This conclusion was further confirmed in the iDPC imaging results with different scan rotation angles (figs. S7 and S8), which indicates that the effect of scan distortion could be neglected.

This geometrical match further reduced the freedom of molecular translation and rotation and constrained confined benzene molecules to adopt nearly identical configurations, where the C_6 planes were parallel to D_{max}. This local deformation of zeolite opening pores suggested that the zeolite framework would change as needed to adapt to the entrance or exit of guest molecules, which is the core of the diffusion or formation of larger guest molecules breaking through the crystallographic pore size limit. Despite such severe local deformation, the overall crystal structure of ZSM-5 zeolite remained intact, and the unit cell parameters were almost unchanged, mainly due to the symmetry of the MFI framework (see below).

We achieved better imaging of O atoms in the zeolite framework during benzene adsorption. Given the 40-nm thickness of the specimen, thermal diffusion and multiple electron scattering could not be ignored and inevitably hampered the imaging resolution (*43*, *44*). Thus, we could only observe Si atoms on the zeolite framework before benzene adsorption, and the O atoms were blurred by nearby Si atoms (Fig. 1H). However, after benzene adsorption, we observed framework O atoms (Fig. 1I). The observed O atoms are highlighted with orange circles in the corresponding crystal structure schematic (Fig. 1J); Fig. 1, H and I, was cropped from fig. S9. We speculated that this difference in O imaging occurred for two reasons. When interacting with confined benzene molecules, the flexible framework was tightened, and the vibration of O atoms in the zeolite framework was restrained, especially in the Si_{10} rings. Moreover, the straight channels were stretched along D_{max} and compressed along D_{min}, which brought an increased distance between adjacent Si-O-Si from the [010] projection.

Dynamic imaging of the benzene desorption process

To further elucidate the host-guest interaction between guest molecules and the zeolite framework, we performed an in situ desorption experiment to investigate the dynamic evolution process of zeolite channels. Real-time iDPC-STEM imaging results at different stages are shown in Fig. 2. Before benzene adsorption, the specimen was also degassed as described above. The holder temperature was decreased to 473 K and stabilized for 30 min. We then introduced the gas mixture (10% benzene and 90% N_2, 450 torr). After 1 min, which was shorter than the time required to obtain a single iDPC-STEM image, the zeolite framework was almost completely filled with benzene (fig. S10). Thus, it was difficult to observe the dynamic details of the adsorption process. In contrast, the desorption of benzene molecules

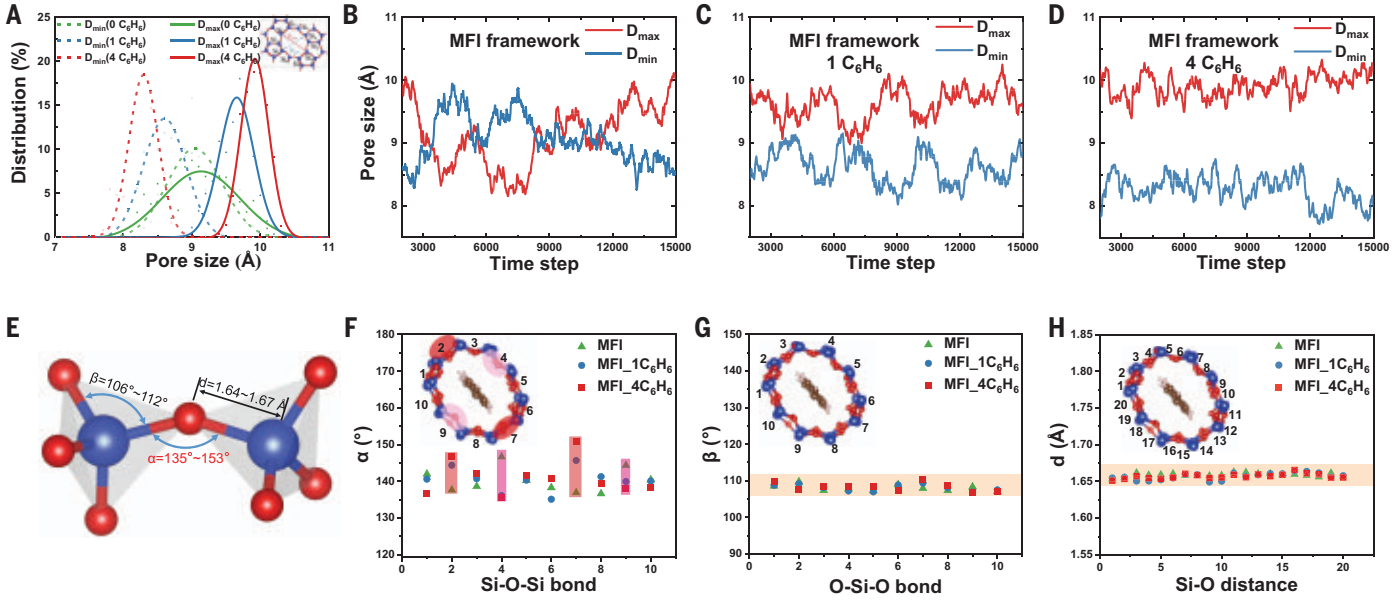


Fig. 3. Investigating the chemical nature of MFI framework flexibility via ab initio molecular dynamics simulation. (**A**) Distributions of pore size for benzene (C_6H_6) molecule adsorption with different loadings as calculated from AIMD simulations at 473 K. (**B** to **D**) Pore size evolution with time for 0 (B), 1 (C), and 4 (D) C_6H_6 molecules adsorbed inside MFI zeolite. (**E**) Schematic diagram of the tetrahedral SiO_4 linkage with three structural defining parameters. (**F** to **H**) Statistical results for the Si-O-Si angle (F), O-Si-O angle (G), and Si-O bond length (H) indicating a loose network composed of rigid tetrahedral SiO_4 and soft hinges at the oxygen atoms. In the statistics of each parameter, the corresponding number of variables is listed in the inserted schematic. See figs. S14 to S17 for details. Time step: 0.5 fs.

was a slow process that could be observed with iDPC-STEM imaging. To quantitatively describe the benzene desorption process, we normalized the contrast of benzene spots in different images according to the contrast of the zeolite framework in the related profile analysis, given that we had been observing the same area (fig. S11).

We recorded the time at which the zeolite framework reached adsorption saturation as time zero; at the same time, we kept the temperature constant and switched the atmosphere to vacuum. Figure 2, A to F, shows real-time images of straight channels throughout the desorption process. As benzene was released, the contrast of the benzene spots in straight channels continuously decreased (initially fast and then more slowly). Most benzene molecules were desorbed within the first 40 min (Fig. 2H). Finally, most of benzene molecules were desorbed after heating to 923 K. Furthermore, we compared the in situ STEM results with the thermal gravimetric analysis (TGA) results (fig. S12). The initial decline rate of the TGA desorption process was faster because of the contribution of benzene molecules adsorbed on the outer surface. Finally, most of the benzene was desorbed in a similar time.

As benzene was released, the channel geometry changed with the number of adsorbed benzene molecules. With the decrease in the contrast of benzene spots, the degree of channel deformation (described as the aspect ratio of D_{max}/D_{min}) gradually decreased, and the shape returned from an ellipse to an approximately round shape (Fig. 2G). The deformation degree of the straight channels was positively correlated with the number of adsorbed benzene molecules in the straight channels; this was further confirmed in saturated ZSM-5 zeolites at different temperatures (fig. S13). This relation indicated that the host-guest interaction was related to the number of adsorbed molecules and would further affect the shape of pores.

Conversely, the geometry of the opening pores also affected the arrangement and orientations of confined benzene molecules (Fig. 2I). With benzene desorption, whereas the pore shapes tended to resemble ellipses with a smaller curvature, the dominant configuration of benzene molecules changed in response (from 52° to 46°) and the angle distribution became wider, indicating a weakened host-guest interaction or confinement to facilitate molecular vibration and rotation. Thus, as the adsorbed benzene molecules deformed the opening pores, the deformed geometry of the zeolite pores dictated the dominant configurations of guest molecules in return. In this regard, this observed alterable geometry of zeolite channels indicated that host-guest interactions in flexible frameworks are not immutable. To better understand the different predominant pathways at different loadings, apart from the interaction between guest molecules, the variation in geometry of the opening pores needs to be considered when we describe the entropy and enthalpy of adsorption (*45*, *46*).

Ab initio molecular dynamics simulations at different loadings

To gain more insight into the subcell deformation of the ZSM-5 zeolites, we performed ab initio molecular dynamics (AIMD) simulations and calculated the distributions of the corresponding lengths of both axes (Fig. 3, A to D). AIMD simulation is a powerful tool for in situ tracking of the dynamic changes of chemical bonds (such as bond lengths and angles) for zeolite catalysts at the picosecond scale under working conditions (*9*, *25*). Both the maximal (D_{max}) and minimal (D_{min}) pore sizes of straight channels (Fig. 3A) with adsorbed benzene molecules of various loadings were detected at the experimental temperature (473 K). The STEM imaging mode obtained images by scanning point by point, so it often took several minutes to form an image and much longer than the time scale of the atomic thermal motion in the framework. However, AIMD simulations can provide more detailed information on zeolite framework thermal motion at a picosecond time scale.

In the absence of adsorbed benzene molecules, the distributions of D_{max} and D_{min} were similar, with peaks located at ~9.1 and ~9.0 Å, respectively. Tracing the variances in pore size over time revealed that the pore channels were highly flexible and tended to be circular statistically, but that there were alternating shifts of D_{max} and D_{min} (Fig. 3B) suggestive of a thermally activated breathing motion of the zeolite pores.

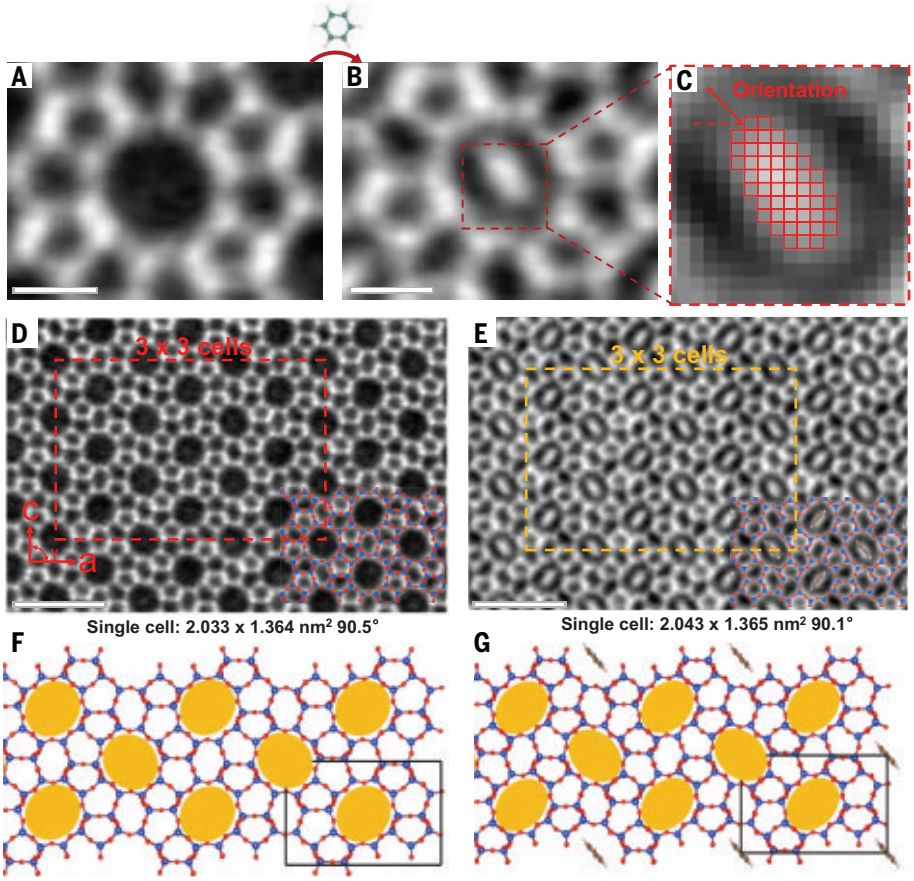


Fig. 4. *Pnma* space group symmetry of MFI zeolite brings subcell topological flexibility and overall rigidity. (**A** to **C**) IDPC images of a straight channel before and after benzene adsorption show severe local deformation, with a high aspect ratio of 1.17. (**D** and **E**) iDPC images of the MFI framework before and after benzene adsorption from the [010] projection with nearly identical cell parameters, from the statistical data of 10 × 10 cells. (**F** and **G**) Corresponding structures of the MFI framework before (F) and after (G) benzene adsorption in [010] projection. The yellow ellipses present the shapes of straight channels. Scale bar, 500 pm [(A) and (B)], 2 nm [(D) and (E)].

However, as more benzene was adsorbed, the flexibility of the pore channels decreased, and the shapes tended to resemble an ellipse with a greater curvature (Fig. 3, C and D). For example, when one benzene molecule was adsorbed per unit cell, the mean D_{max} and D_{min} were 9.6 and 8.6 Å, respectively. The deviation between D_{max} and D_{min} increased (to ~9.9 and 8.3 Å, respectively) when four benzene molecules were confined inside each MFI zeolite unit cell. Moreover, the standard deviation of the pore size distribution tended to decrease as more benzene adsorbed, indicating decreasing flexibility in the Si_{10} opening pores (Fig. 3A and table S1). Overall, as the amount of adsorbed benzene increased, the deformation of MFI straight channels increased, which was consistent with our experimental images.

Structurally, zeolites are composed of corner-sharing nearly perfect tetrahedral TO_4, connected by hinges at the oxygen atoms (fig. S1). Hence, the different structures of zeolites are mainly determined by three structural parameters: α (Si-O-Si bond angle), β (O-Si-O bond angle), and *d* (Si-O bond length) (Fig. 3E). To further investigate the origin of MFI zeolite flexibility, we performed detailed statistical analyses on the bond lengths and angles in the 10-membered rings (10MRs). With benzene adsorption, the main contribution to channel deformation came from the increase in α (Fig. 3F and fig. S15), whereas β and *d* within tetrahedral SiO_4 were almost unchanged (Fig. 3, G and H, and figs. S16 and S17). The most substantial change came from the second and seventh α, which were exactly vertical to the plane of the benzene molecule (Fig. 3F). The local stress was mainly concentrated in these two Si-O-Si bonds and nearby areas, causing the Si-O-Si angle to be flattened and distant in the [010] projection (fig. S18).

Previous studies have found that thermodynamics and kinetics are sensitive to the difference in the local chemical environment, especially the T-O-T angle at the acid site, which can be stretched to influence the intrinsic acidity, stabilize the protonated transition state, and thereby control the reaction pathway in ZSM-5 (*47*, *48*), which would have a profound impact on understanding the molecular reaction mechanism. In contrast, the distributions of β and *d* at different loadings were nearly identical and near the theoretical value of a perfect SiO_4 tetrahedron (109° and 1.66 Å) (Fig. 3, G and H). We further examined the distributions of six O-Si-O angles within a single tetrahedral SiO_4 in a 10MR straight channel to confirm the rigidity of the SiO_4 tetrahedron (fig. S19). Thus, we concluded that the local flexibility of the zeolite framework originated from the topologically soft Si-O-Si hinges between adjacent tetrahedral SiO_4, whereas the tetrahedral SiO_4 remained rigid and nearly perfect.

MFI zeolite framework with subcell topological flexibility and overall rigidity

The above structural analysis and theoretical calculation results revealed the local flexibility in the MFI framework, which allows extensive local deformation of the straight channels to accommodate larger guest molecules. When benzene and other organic sorbates, with a comparable molecular size to the opening pores, adsorb from the gas phase into the pore-confined phase, the molecules fall into a deep potential well. The movement and vibrations of confined molecules are restricted, and the images of confined molecules resemble their molecular dimensions, behaving as one-dimensional oriented single crystals, such that the molecules cannot simply be treated as spheres. Thus, the opening pores will be stretched according to the molecular dimensions of the confined molecules, and the maximum change in pore diameter can reach 0.63 Å in the [010] projection (Fig. 4, A to C, and figs. S20 and S21).

Despite severe subcell deformation, monocrystalline MFI zeolite could still maintain an intact crystal structure, and the unit cell parameters from the [010] projection were basically unchanged (Fig. 4, D and E). Here, we used profile analysis to study the unit cell parameters (*a*, *c*) and their included angle from the [010] projection. We only saw slight variance in the overall cell parameters (from 2.033 ± 0.002 nm × 1.364 ± 0.001 nm × 90.5° to 2.043 ± 0.002 nm × 1.365 ± 0.001 nm × 90.1°). The changes in cell size and included angle were < 0.5%, indicating little structural change in the bulk material, which was consistent with previous work (*18*, *20*, *23*).

This considerable difference between local and global structural changes may originate in the symmetry of the zeolite framework. Although part of the TO_4 tetrahedrons was twisted by host-guest interactions to form local deformation in opening pores, the *Pnma* space group symmetry of the MFI framework makes the subcell deformation of adjacent straight channels symmetrical, and the changes in D_{max} and D_{min} appear to alternately compensate for each other (Fig. 4, F and G). Thus, the

particular symmetry of the MFI framework gives it local flexibility and overall rigidity, which allows subcell deformation to accommodate larger molecules but still maintain the stability of the entire zeolite framework. We termed this structural characteristic of the MFI zeolite as subcell topological flexibility.

Summary

Combining iDPC-STEM imaging with an atmosphere system, we achieved in situ observation of the dynamic subcell topological deformation of MFI channels with atomic resolution while the whole MFI zeolite single-crystal structure was maintained nearly unchanged. We revealed the intrinsic mechanism of a molecule breaking through the pore size limit. Upon benzene adsorption, the straight channels underwent severe local deformation along the unified orientations of guest molecules, which allowed the diffusion of larger molecules.

This subcell topological flexibility mainly comes from the soft Si-O-Si hinges between rigid tetrahedral SiO_4, verified in AIMD simulations and our experimental results. In the in situ benzene desorption process, we captured the dynamic evolution of zeolite channels with decreasing numbers of benzene molecules, indicating that the degree of subcell local deformation is positively correlated with the shape and amount of confined benzene molecules. Furthermore, the symmetry of the MFI framework enables mutual compensation by large deformation of adjacent channels, thereby keeping the overall crystal structure stable and resulting in macroscopically rigid behavior. As a consequence, the MFI framework is highly industrially applicable in the family of zeolites. These results confirmed the microscopic subcell topological flexibility of a macroscopically rigid zeolite framework and revealed its chemical nature, advancing our understanding of the intrinsic mechanism of molecular diffusion breaking through the pore size limit and the impact of subcell topological flexibility on adsorption and catalytic transformation in microporous materials.

REFERENCES AND NOTES

1. D. W. Breck, *Zeolite Molecular Sieves: Structure, Chemistry, and Use* (Wiley, 1973).
2. M. W. Deem, J. M. Newsam, J. A. Creighton, *J. Am. Chem. Soc.* **114**, 7198–7207 (1992).
3. C. Baerlocher, W. M. Meier, D. H. Olson, *Atlas of Zeolite Framework Types* (Elsevier, 2001).
4. B. Smit, T. L. M. Maesen, *Nature* **451**, 671–678 (2008).
5. E. G. Derouane, in *Studies in Surface Science and Catalysis*, B. Imelik *et al.*, Eds. (Elsevier, 1980), vol. 5, pp. 5–18.
6. E. G. Derouane, Z. Gabelica, *J. Catal.* **65**, 486–489 (1980).
7. T. F. Degnan Jr., *J. Catal.* **216**, 32–46 (2003).
8. Z. Lai *et al.*, *Science* **300**, 456–460 (2003).
9. P. J. Bereciartua *et al.*, *Science* **358**, 1068–1071 (2017).
10. J. Jae *et al.*, *J. Catal.* **279**, 257–268 (2011).
11. S. Horike, S. Shimomura, S. Kitagawa, *Nat. Chem.* **1**, 695–704 (2009).
12. H. Furukawa, K. E. Cordova, M. O'Keeffe, O. M. Yaghi, *Science* **341**, 1230444 (2013).
13. L. Sarkisov, R. L. Martin, M. Haranczyk, B. Smit, *J. Am. Chem. Soc.* **136**, 2228–2231 (2014).
14. Z. Wang, R. F. Lobo, J. Lambros, *Microporous Mesoporous Mater.* **57**, 1–7 (2003).
15. Z. Li *et al.*, *Angew. Chem. Int. Ed.* **45**, 6329–6332 (2006).
16. G. N. Greaves, A. L. Greer, R. S. Lakes, T. Rouxel, *Nat. Mater.* **10**, 823–837 (2011).
17. E. L. Wu, S. L. Lawton, D. H. Olson, A. C. Rohrman, G. T. Kokotailo, *J. Phys. Chem.* **83**, 2777–2781 (1979).
18. D. G. Hay, H. Jaeger, *J. Chem. Soc. Chem. Commun.* **1984**, 1433 (1984).
19. C. A. Fyfe, H. Strobl, G. T. Kokotailo, G. J. Kennedy, G. E. Barlow, *J. Am. Chem. Soc.* **110**, 3373–3380 (1988).
20. W. C. Conner, R. Vincent, P. Man, J. Fraissard, *Catal. Lett.* **4**, 75–83 (1990).
21. C. A. Fyfe, Y. Feng, H. Grondey, G. T. Kokotailo, H. Gies, *Chem. Rev.* **91**, 1525–1543 (1991).
22. S. Quartieri, G. Montagna, R. Arletti, G. Vezzalini, *J. Solid State Chem.* **184**, 1505–1516 (2011).
23. M. Ardit, A. Martucci, G. Cruciani, *J. Phys. Chem. C* **119**, 7351–7359 (2015).
24. A. Sartbaeva, S. A. Wells, M. M. J. Treacy, M. F. Thorpe, *Nat. Mater.* **5**, 962–965 (2006).
25. H. Shi, A. N. Migues, S. M. Auerbach, *Green Chem.* **16**, 875–884 (2014).
26. S. Caro-Ortiz, E. Zuidema, M. Rigutto, D. Dubbeldam, T. J. H. Vlugt, *J. Phys. Chem. C* **124**, 24488–24499 (2020).
27. L. Liu *et al.*, *Angew. Chem. Int. Ed.* **59**, 819–825 (2020).
28. B. Shen *et al.*, *Adv. Mater.* **32**, e1906103 (2020).
29. B. Shen *et al.*, *Nature* **592**, 541–544 (2021).
30. B. Shen *et al.*, *Nat. Commun.* **12**, 2212 (2021).
31. I. Lazić, E. G. T. Bosch, S. Lazar, *Ultramicroscopy* **160**, 265–280 (2016).
32. E. Yücelen, I. Lazić, E. G. T. Bosch, *Sci. Rep.* **8**, 2676 (2018).
33. M. Choi *et al.*, *Nature* **461**, 246–249 (2009).
34. Y. Liu *et al.*, *ChemCatChem* **5**, 1517–1523 (2013).
35. Z. Shan *et al.*, *Chem. Commun.* **47**, 1048–1050 (2011).
36. Y. Ma *et al.*, *RSC Advances* **6**, 81198–81202 (2016).
37. J. Yang *et al.*, *J. Energy Chem.* **35**, 44–48 (2019).
38. P. Payra, P. K. Dutta, S. Auerbach, K. Carrado, *Handbook of Zeolite Science and Technology* (Routledge, 2003).
39. X. Yue, X. Yang, *Langmuir* **22**, 3138–3147 (2006).
40. O. Gobin, S. Reitmeier, A. Jentys, J. Lercher, *Microporous Mesoporous Mater.* **125**, 3–10 (2009).
41. B. Smit, T. L. M. Maesen, *Nature* **374**, 42–44 (1995).
42. V. R. Choudhary, V. S. Nayak, T. V. Choudhary, *Ind. Eng. Chem. Res.* **36**, 1812–1818 (1997).
43. J. M. Cowley, A. F. Moodie, *Acta Crystallogr.* **10**, 609–619 (1957).
44. S. Hillyard, J. Silcox, *Ultramicroscopy* **58**, 6–17 (1995).
45. K. Alexopoulos *et al.*, *J. Phys. Chem. C* **120**, 7172–7182 (2016).
46. D. Mei, J. A. Lercher, *AIChE J.* **63**, 172–184 (2017).
47. J. A. Ryder, A. K. Chakraborty, A. T. Bell, *J. Phys. Chem. B* **104**, 6998–7011 (2000).
48. J. Dědeček *et al.*, *J. Phys. Chem. C* **113**, 1447–1458 (2009).

ACKNOWLEDGMENTS

We thank F. Lu and B. Y. Shen for discussions, and Tsinghua National Laboratory for Information Science and Technology for assistance with the energy simulation. **Funding:** National Key Research and Development Program of China grants 2020YFB0606401, 2018YFE0122600, 2018YFB0604801, and 2021YFA1502600; National Natural Science Foundation of China grants 22005170, 22125304, and 22032005. **Author contributions:** Conceptualization: F.W., X.C. Methodology: H.X., X.C., Z.L., H.W., Investigation: H.X., X.C., Z.L. Visualization: H.X., X.C., Z.L. Supervision: F.W., X.C., C.Z., A.Z. Writing–original draft: H.X. Writing–review and editing: F.W., X.C., C.Z., H.X., W.Q., Z.L., A.Z. **Competing interests:** Authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and the supplementary materials.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abn7667
Materials and Methods
Figs. S1 to S21
Table S1
Movie S1
Data S1 to S3
References (*49*–*53*)

17 December 2021; accepted 16 February 2022
10.1126/science.abn7667

REPORTS

DYNAMIC GENOME

Dynamics of CTCF- and cohesin-mediated chromatin looping revealed by live-cell imaging

Michele Gabriele[1,2,3]†, Hugo B. Brandão[1,2,3]†‡, Simon Grosse-Holz[4,5]†, Asmita Jha[1,2,3], Gina M. Dailey[6], Claudia Cattoglio[6,7], Tsung-Han S. Hsieh[6,7], Leonid Mirny[4,5,8]*, Christoph Zechner[9,10,11,12]*, Anders S. Hansen[1,2,3]*

Animal genomes are folded into loops and topologically associating domains (TADs) by CTCF and loop-extruding cohesins, but the live dynamics of loop formation and stability remain unknown. Here, we directly visualized chromatin looping at the *Fbn2* TAD in mouse embryonic stem cells using super-resolution live-cell imaging and quantified looping dynamics by Bayesian inference. Unexpectedly, the *Fbn2* loop was both rare and dynamic, with a looped fraction of approximately 3 to 6.5% and a median loop lifetime of approximately 10 to 30 minutes. Our results establish that the *Fbn2* TAD is highly dynamic, and about 92% of the time, cohesin-extruded loops exist within the TAD without bridging both CTCF boundaries. This suggests that single CTCF boundaries, rather than the fully CTCF-CTCF looped state, may be the primary regulators of functional interactions.

Mammalian genomes are folded into loops and domains known as topologically associating domains (TADs) by the proteins CTCF and cohesin (*1*). Mechanistically, cohesin is thought to load on DNA and bidirectionally extrude loops until it is blocked by CTCF such that CTCF establishes TAD boundaries (*2*–*6*). Functionally, CTCF- and cohesin-mediated looping and TADs play critical roles in multiple nuclear processes, including regulation of gene expression, somatic recombination, and DNA repair

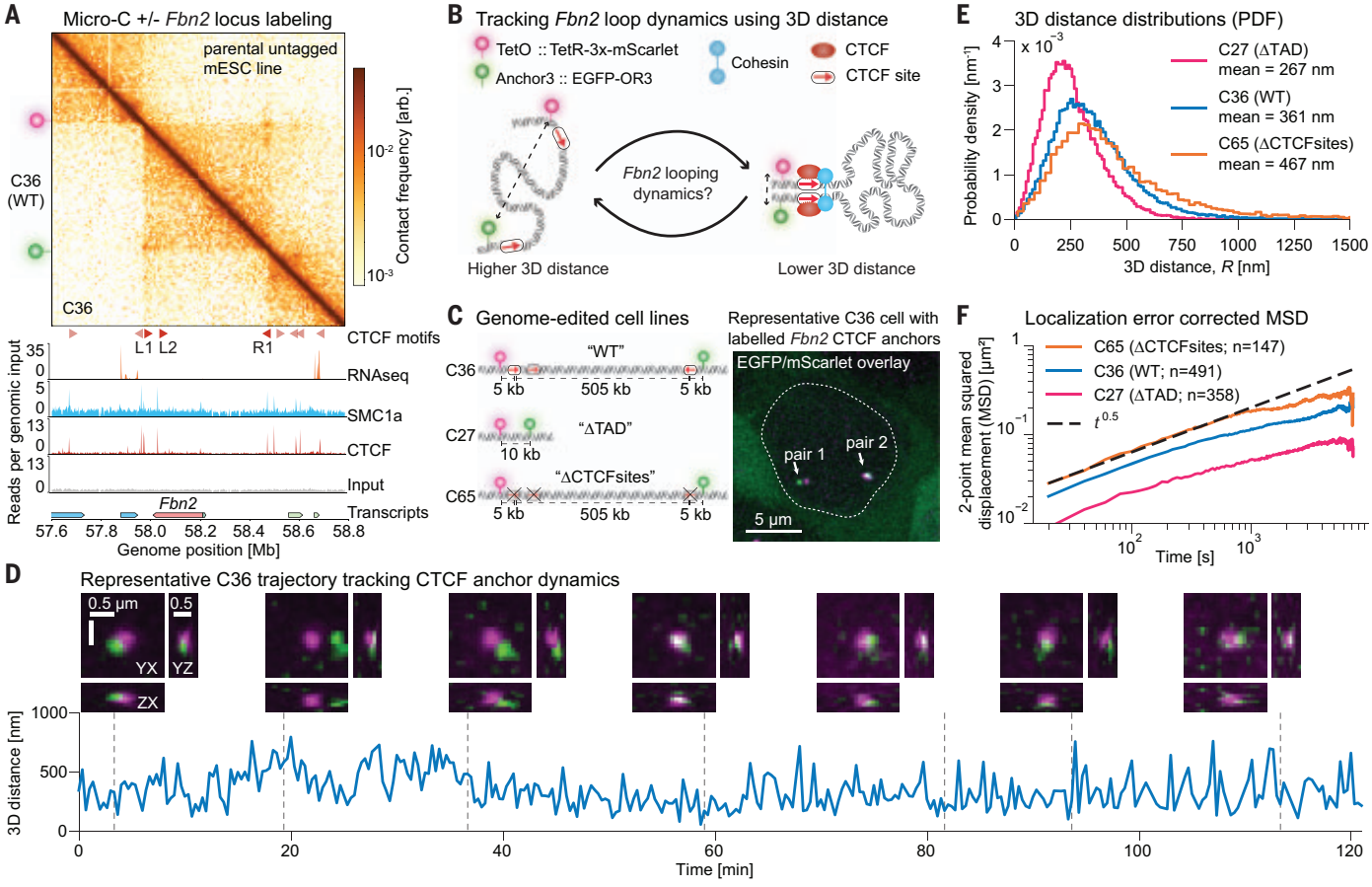


Fig. 1. Endogenous labeling and tracking of the *Fbn2* loop with super-resolution live-cell imaging. (**A**) Fluorescent labeling of *Fbn2* loop anchors does not perturb the *Fbn2* TAD. Shown is a Micro-C contact map comparing the parental untagged (C59, top left) and tagged (C36, bottom right) cell lines. Red triangles are CTCF motifs with orientation. C36 ChIP-seq shows CTCF (GSM3508478) and cohesin (SMC1A; GSM3508477) binding compared with input (GSM3508475). *Fbn2* is not expressed (RNA-seq GSE123636; annotation: GRCm38). Genome coordinates: mm10. (**B**) Overview of tagging and readout using 3D distance. (**C**) Overview of the genome-edited cell lines (left) and a representative maximum intensity projection (MIP) of a cell nucleus showing two pairs of "dots" (right). (**D**) Representative 3D trajectory over time of a dot pair. MIPs of the 3D voxels centered on the mScarlet dot (top) and 3D distances between dots (bottom) are shown. (**E**) 3D distance probability density functions of dot pairs (n = 32,171, n = 46,163, and n = 13,566 distance measurements for C27, C36, and C65, respectively). (**F**) Localization error–corrected two-point MSD plots (n = 358, n = 491, and n = 147 trajectories in C27, C36, and C65, respectively).

(*1*). For example, TADs are thought to regulate gene expression by increasing the frequency of enhancer-promoter interactions within a TAD and by decreasing enhancer-promoter interactions between TADs (*7*). However, to understand how TADs and loops are formed and maintained and how they function, it is necessary to understand the dynamics of CTCF- and cohesin-mediated loop formation and loop lifetime.

Although recent advances in single-cell genomics and fixed-cell imaging have made it possible to generate static snapshots of three-dimensional (3D) genome structures in single cells (*8*–*13*), live-cell imaging is required to understand the dynamics of chromatin looping (*14*). Furthermore, previous studies have yielded conflicting results as to whether loops are well defined in single cells (*8*–*13*), perhaps because of the difficulty associated with distinguishing bona fide CTCF- and cohesin-mediated loops from mere proximity that emerges stochastically (*14*). Recent pioneering work has visualized enhancer-promoter interactions (*15*–*17*) and long-range V(D)J–chromatin interactions (*18*) in live cells. However, the dynamics of loop formation and the lifetime of CTCF- and cohesin-mediated loops have not yet been quantified in living cells.

To visualize the dynamics of CTCF- and cohesin-mediated looping, we chose as our model system the loop holding together the two CTCF-bound boundaries of the 505-kb *Fbn2* TAD in mouse embryonic stem cells (mESCs). This TAD is verified to be CTCF dependent (*19*) and relatively simple because it only contains a single gene, *Fbn2*, which is not expressed in mESCs (Fig. 1A). We used genome editing to homozygously label the left and right CTCF sites of the *Fbn2* TAD with TetO and Anchor3 arrays, which we then visualized by coexpressing the fluorescently tagged binding proteins TetR-3x-mScarlet and EGFP-OR3 (*20*) [clone C36 (Fig. 1, B to D)]. We developed a comprehensive image analysis framework, *ConnectTheDots*, to extract trajectories of 3D loop anchor positions from the acquired movies (fig. S1). By optimizing 3D super-resolution live-cell imaging

[1]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [2]The Broad Institute of MIT and Harvard, Cambridge, MA 02139, USA. [3]Koch Institute for Integrative Cancer Research, Cambridge, MA 02139, USA. [4]Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [5]Institut Curie, 75005 Paris, France. [6]Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, CA 94720, USA. [7]Howard Hughes Medical Institute, University of California, Berkeley, Berkeley, CA 94720, USA. [8]Institute for Medical Engineering and Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [9]Max Planck Institute of Molecular Cell Biology and Genetics, 01307 Dresden, Germany. [10]Center for Systems Biology Dresden, 01307 Dresden, Germany. [11]Cluster of Excellence Physics of Life, TU Dresden, 01307 Dresden, Germany. [12]Faculty of Computer Science, TU Dresden, 01187 Dresden, Germany.
***Corresponding author. Email: leonid@mit.edu (L.M.); zechner@mpi-cbg.de (C.Z.); ashansen@mit.edu (A.S.H.)**
†These authors contributed equally to this work.
‡Present address: Illumina Inc., San Diego, CA 92122, USA.

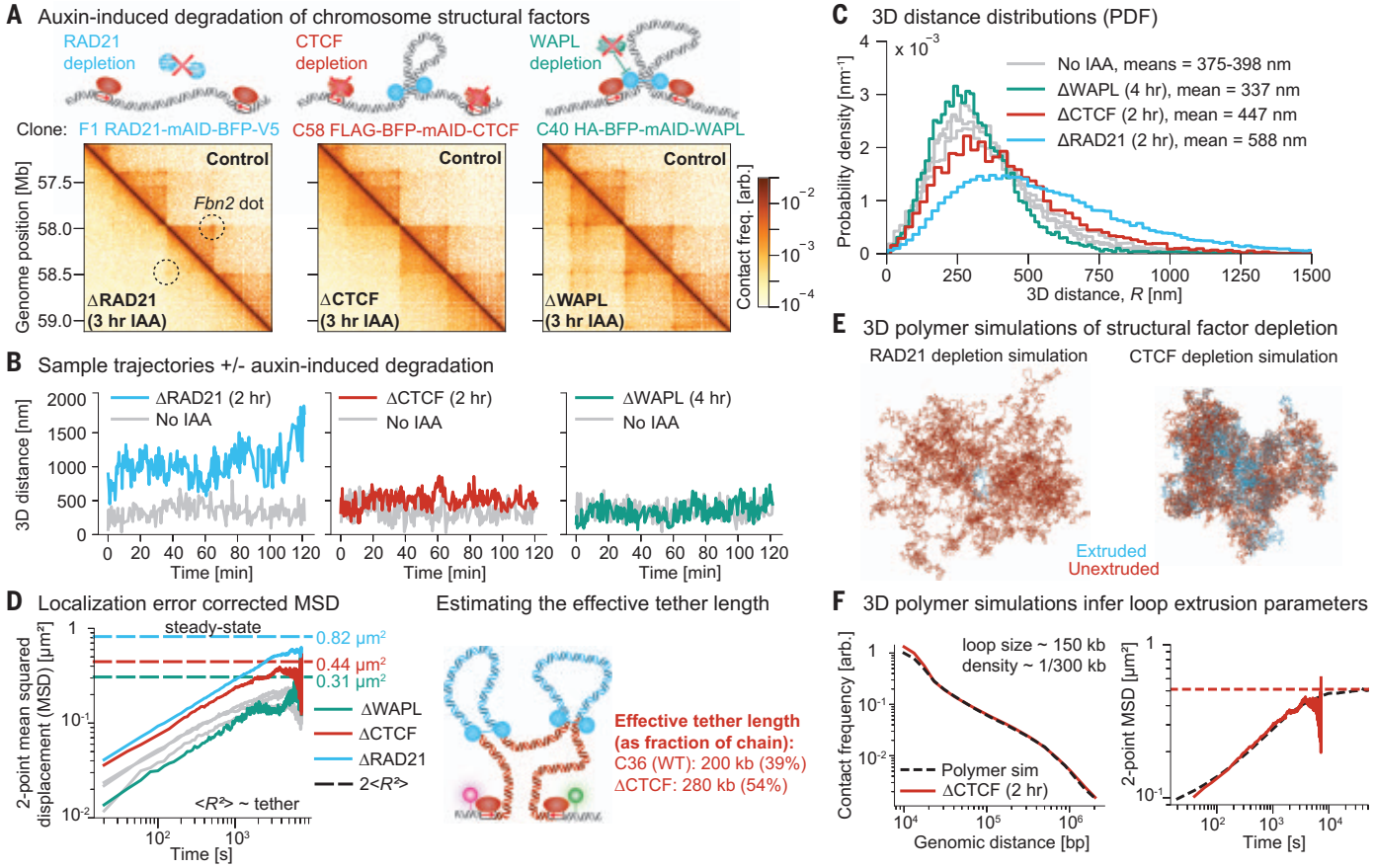


Fig. 2. Degradation of CTCF, cohesin, and WAPL reveals their role in loop extrusion and looping-mediated chromosome compaction. (**A**) Micro-C data for the AID-tagged clones for RAD21 (left), CTCF (middle), and WAPL (right) showing control data [no indole-3-acetic acid (IAA) treatment; top half] and protein degradation data (3 hours after IAA; bottom half), with schematics illustrating the expected effect. (**B**) Representative trajectories with (colored lines) or without (gray lines) IAA treatment for each AID-tagged clone. (**C**) 3D distance probability density functions of dot pairs [n = 45,379, n = 10,469, and n = 18,153 distance measurements for ΔRAD21 (2 hours), ΔCTCF (2 hours), and ΔWAPL (4 hours) depletion conditions, respectively, and n = 17,605, n = 11,631, and n = 21,001 for the same clones without treatment]. (**D**) Localization error–corrected two-point MSD plots for the AID-tagged clones (left) [n = 537, n = 137, and n = 215 trajectories in ΔRAD21 (2 hours), ΔCTCF (2 hours), and ΔWAPL (4 hours) depletion conditions, respectively, and n = 183, n = 151, and n = 257 for the same clones without treatment (gray lines)]. The effective tether length was obtained by computing the ratio of the steady-state variance of each clone to the value in the RAD21-depletion condition (note that $2\langle R^2\rangle$ is also the asymptotic value of the MSD; see also the supplementary materials). (**E**) Representative 3D polymer conformation from simulations mimicking the ΔRAD21 (95% cohesin depletion) (left) and ΔCTCF (100% CTCF depletion) (right) depletion conditions. Red is the unextruded segment, blue the extruded segment. (**F**) Matching simulations to the data to obtain loop-extrusion parameters (three fit parameters; see the supplementary materials).

conditions (*14*), we could track *Fbn2* looping dynamics at 20-s resolution for >2 hours (Fig. 1D and movies S1 to S4). After DNA replication in the S/G_2 phase, it is no longer possible to reliably distinguish intrachromosomal from sister-chromosomal interactions (*14*). We therefore filtered out replicated and low-quality dots using a convolutional neural network (fig. S2). Thus, we only considered G_1 and early S-phase cells.

To validate our system for tracking *Fbn2* loop dynamics, we performed a series of control experiments. First, we confirmed using Micro-C (*21*, *22*) that our locus-labeling approach did not measurably perturb the *Fbn2* loop (Fig. 1A). Second, to "mimic" the looped state, we deleted the 505 kb between the CTCF sites, generating clone C27 ("ΔTAD"; Fig. 1C). As expected, this reduced the 3D distance (Fig. 1E; the nonzero 3D distance distribution for C27 was expected because of localization noise and the 5-kb tether between CTCF sites and fluorescent labels; see fig. S3). Third, as a negative control for CTCF-mediated looping, we generated clone C65 ("ΔCTCFsites"; Fig. 1C) by homozygously deleting the three CTCF motifs in the *Fbn2* TAD (L1, L2, R1; Fig. 1A) and validated that this resulted in a loss of CTCF binding and cohesin colocalization using chromatin immunoprecipitation sequencing [ChIP-seq (fig. S4)]. As expected, the 3D distance increased in C65 (Fig. 1E). Next, we calculated mean-squared displacements (MSDs) of the relative positions of the two loci (two-point MSD), which are unaffected by cell movement. Chromatin dynamics were consistent with Rouse polymer dynamics, with a scaling of MSD ~ $t^{0.5}$ for all three clones (*23*) (Fig. 1F). We conclude that our approach faithfully reports on CTCF looping dynamics in live cells without noticeable artifacts.

To elucidate the specific roles of CTCF and cohesin, we endogenously tagged RAD21, CTCF, and the cohesin unloader WAPL with mAID in the C36 (WT) line, allowing for degradation with indole-3-acetic acid (*24*). For RAD21 and CTCF, we achieved near-complete depletion in 2 hours (fig. S5), long-term depletion led to cell death (fig. S6), Micro-C analysis revealed loss of the *Fbn2* loop or corner peak as expected (*25*–*28*) (Fig. 2A), and ChIP-Seq

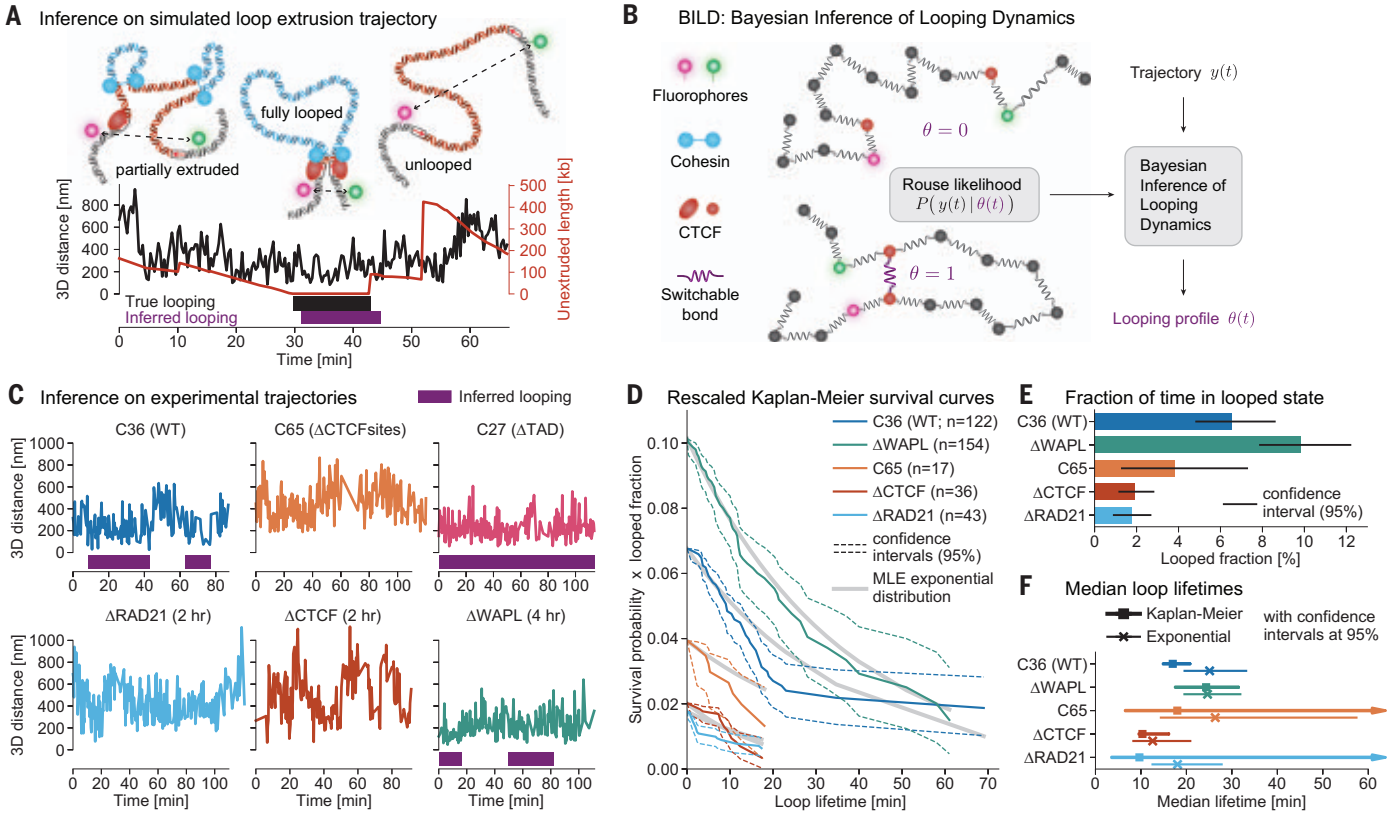


Fig. 3. BILD reveals rare and dynamic CTCF loops. (**A**) Example trajectory from polymer simulations with loop extrusion. Extrusion shortens the effective tether (red is the unextruded length, ground truth from simulations) between the CTCF sites. A ground truth loop is formed when the tether is minimal and cohesin is stalled at both CTCF sites (black bar). BILD captures these accurately (purple bar). (**B**) Schematic overview of BILD. Building on the analytical solution to the Rouse model, we used a hierarchical Bayesian model to determine the optimal looping profile for single trajectories. (**C**) Illustrative examples of inferred looping on real trajectory data. (**D**) Kaplan-Meier survival curves rescaled by the inferred looped fraction. Gray lines are maximum likelihood fits of a single exponential to the data, accounting for censoring. (**E**) Fraction of time the *Fbn2* locus spends in the fully looped conformation. Error bars are bootstrapped 95% confidence intervals. (**F**) Median loop lifetimes from the Kaplan-Meier survival curves (squares) or exponential fits (crosses). Confidence intervals are determined from the confidence intervals on the Kaplan-Meier curve and the likelihood function of the exponential fit, respectively. Where the upper confidence limit on the survival curve did not cross below 50%, an arrowhead indicates a semi-infinite confidence interval.

analysis showed a loss of chromatin binding (fig. S7). For WAPL, depletion took 4 hours and was less complete (fig. S5), long-term depletion occasionally yielded visibly compacted "vermicelli" chromosomes (*29*) (fig. S6C), and Micro-C analysis revealed increased corner peak strength (*27*, *28*, *30*) (Fig. 2A). All three AID lines exhibited lower protein abundance, likely because of leaky protein depletion (fig. S8).

Having validated the AID cell lines, we next performed live-cell imaging to study the specific roles of RAD21, CTCF, and WAPL in loop extrusion in vivo. Consistent with RAD21 being required for loop extrusion, RAD21 depletion strongly increased the 3D distances (Fig. 2, B and C). Consistent with CTCF being the boundary factor that is required for *Fbn2* loop formation (Fig. 1B) but not required for loop extrusion, CTCF depletion increased 3D distances, albeit less than RAD21 depletion (*6*). Finally, consistent with prior observations that WAPL depletion increases cohesin residence time and abundance on chromatin (*29*), potentially allowing it to extrude longer and more stable loops (*27*, *30*), WAPL depletion decreased the 3D distances (Fig. 2, B and C).

To quantify the extent of loop extrusion of the *Fbn2* TAD, we turned to polymer physics theory. The Rouse model predicts a linear relationship between chain length and mean squared distance ($<R^2>$) between the fluorescent labels (dashed lines in Fig. 2D; see also fig. S9). By assuming that ΔRAD21 represents the fully unextruded state with a genomic separation of 515 kb (Fig. 1C), we could then assign an "effective tether length" (i.e., the unextruded fraction) to each condition. We found an effective tether length of ~200 kb in C36 (WT) and ~280 kb in ΔCTCF, corresponding to ~39 and ~54% of the full genomic separation, respectively. By subtraction, the genomic separation between the two labels shortened by ~46% due to extrusion alone (ΔRAD21 versus ΔCTCF) and by ~61% due to extrusion with boundaries (ΔRAD21 versus C36). This shows that by blocking extruding cohesins, CTCF increases the fraction extruded between the two CTCF boundaries. Overall, we estimate that, on average, just over half of the *Fbn2* TAD is extruded.

By combining these measurements (Fig. 2, B to D) with our Micro-C data (Fig. 2A), we were then able to determine dynamic parameters of our polymer model of loop extrusion (Fig. 2, E and F), including spacing between cohesins and their processivity, as well as the total strength of the CTCF boundaries (see the supplementary materials). Consistent with our ΔRAD21 data, our polymer simulations resulted in chromosome decompaction after near-complete RAD21 depletion (Fig. 2E) and accurately matched our experimental data (Fig. 2F).

Next, we sought to identify where and when CTCF-CTCF loops occured in our trajectories.

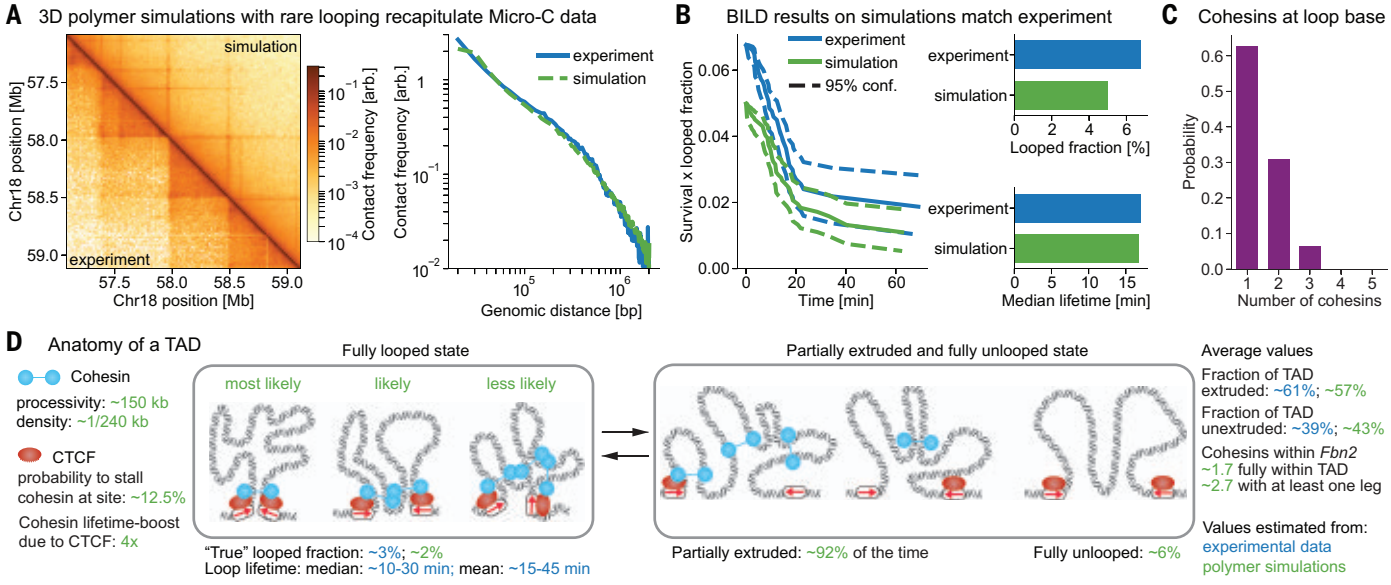


Fig. 4. Comprehensive picture of the *Fbn2* TAD. (**A**) Comparison of Micro-C data for the C36 (WT) line with in silico Micro-C of our best-fit simulation map (left) and contact probability scaling (right). (**B**) BILD applied to the same simulation (green) compared with C36 (WT) experimental data (blue). (**C**) Number of cohesins forming the looped state in simulations (n = 18,789). (**D**) "Anatomy" of the *Fbn2* TAD. Quantitative description of the *Fbn2* TAD is shown using both real data (blue) and our best-fit simulation (green). Cohesin processivity and density and CTCF stalling probability and lifetime boost are simulation parameters. Fraction of time in different conformations was extracted from simulation ground truth using effective tether lengths of 1.1 and 505 kb as cutoffs to define "fully looped" and "fully unlooped," respectively. The fraction of TAD that was unextruded was calculated using the mean tether length over the full simulation.

Because of localization noise and substantial temporal correlations in the data, simple analysis methods failed when benchmarked on simulations (see the supplementary text). Therefore, we developed Bayesian inference of looping dynamics (BILD). In BILD, we coarse-grained the possible conformations of the TAD (Fig. 3A) into two states: (i) a state of sustained contact between the CTCF sites, presumably mediated by cohesin (the "looped state"), and (ii) all other possible conformations, including partial extrusion, random contacts, and the fully unlooped conformation (the "unlooped state"). While the looped state relies on CTCF activity, the unlooped state reflects extrusion without bridged CTCF boundaries, resembling the ΔCTCF condition. On the basis of the MSD~$t^{0.5}$ scaling observed in Fig. 2D, we modeled the unlooped state as a free Rouse chain calibrated to the ΔCTCF data (fig. S10). To model the looped state, we introduced an additional bond between the two CTCF sites (Fig. 3B); this bond is switchable, allowing transitions between the looped and unlooped states. The length of the bond was set to reproduce the 10-kb distance between the fluorophores, using ΔRAD21 as reference for a free 515-kb chain (see the supplementary text). Finally, by using a hierarchical Bayesian model (*31*), BILD then exploited the different spatiotemporal dynamics of the looped state to infer which segments of each trajectory were in the looped state (purple segment in Fig. 3A). When tested on 3D polymer simulations with experimentally realistic noise, BILD accurately inferred both the looped fraction and the loop lifetime (Fig. 2, E and F, and figs. S11 and S12). In summary, BILD allows us to distinguish CTCF- and cohesin-mediated looping from mere proximity.

We next used BILD to infer looping in our experimental trajectory data (Fig. 3, C to F). BILD revealed that the *Fbn2* TAD was fully looped ~6.5% (~3%) of the time, but spent ~93.5% (97%) of the time in a fully unlooped or partially extruded conformation (Fig. 3E). We use brackets to indicate the looped fraction after false-positive correction (fig. S12; the corrected looped fraction was ~6% when we calibrated BILD using a 15-kb fluorophore distance; fig. S13). By contrast, we observed a minimal looped fraction of ~2% (~0%) in ΔRAD21 and ΔCTCF and ~4% (~1%) in C65 (ΔCTCFsites), whereas the looped fraction was increased to ~10% (~6%) in ΔWAPL, consistent with WAPL unloading cohesin from chromatin (*29*).

Finally, we estimated the lifetime of the looped state (Fig. 3, D and F). Accurate measurement of loop lifetimes from finite trajectories can be challenging when trajectories begin or end in the looped state, so that it is unclear how long the looped period truly lasted (e.g., the looped state in the ΔWAPL trajectory in Fig. 3D existed an unknown time before the start of the movie). This problem, known in medical statistics as "censoring," can be solved using the Kaplan-Meier survival estimator. Using this approach, we obtained censoring-corrected survival curves (Fig. 3D) of the looped state, from which we estimated the median loop lifetime (Fig. 3F). Orthogonal to this non-parametric analysis, we also fitted an exponential model, yielding similar estimates. Together, these results give an estimate of the median loop lifetime of ~10 to 30 min in C36 (WT) (Fig. 3F and fig. S12D). These results revealed the fully looped state to be both rare (~3 to 6%) and dynamic (median ~10 to 30 min; mean ~15 to 45 min). Thus, during an average ~12-hour mESC cell cycle, the looped state will occur approximately one or two times, lasting cumulatively ~20 to 45 min, but the remaining ~11.5 hours will be in the partially extruded or fully unlooped conformations.

To understand whether a low looped fraction of ~3% is consistent with a clear and strong corner peak in the Micro-C map, we set up polymer simulations with loop extrusion. Consistent with recent reports, we found that CTCF-mediated stabilization of cohesin was necessary to reproduce both of these features in our simulations (Fig. 4 and fig. S14). We confirmed this effect using inverse fluorescence recovery after photobleaching (iFRAP) of cohesin, finding that CTCF depletion decreased cohesin residence time (fig. S15). Incorporating this effect, we then simulated loop extrusion with a cohesin density of 1/240 kb and processivity of 150 kb (processivity = lifetime × extrusion

speed). When cohesin reaches a CTCF site, it has a 12.5% probability of stalling, which, using the estimate of 50% CTCF occupancy (*34*), translates into an ~25% capture efficiency of CTCF. Once stalled on one side by CTCF, cohesin is stabilized fourfold beyond its base lifetime of ~20 min (*35*) (fig. S14), facilitating the formation of longer loops because the other side of cohesin may continue to extrude. These simulations reproduced both our experimental Micro-C maps (Fig. 4A) and the median loop lifetime and low looped fraction (Fig. 4B).

Together, these results allow us to paint a comprehensive mechanistic picture of the *Fbn2* TAD (Fig. 4, C and D). Most of the time (~92%), the TAD is partially extruded, with ~57 to 61% of the *Fbn2* region captured in one to three extruding cohesin loops, whereas ~39 to 43% remain unextruded. The fully unlooped conformation, as would be found in the absence of cohesin, occurs only ~6% of the time, whereas the fully looped state is even more rare at ~3% (~2% in simulations) and has a median lifetime of ~10 to 30 min. Our simulations reveal that the looped state is sometimes held together by multiple cohesins (Fig. 4C), which also explains why the loop lifetime can be substantially shorter than the CTCF-stabilized cohesin lifetime. We stress that both the mechanistic assumptions of our polymer models and the experimental data constraining them are associated with uncertainty, resulting in uncertainty of the inferred parameters (Fig. 4D). For example, if we allow extruding cohesins to bypass each other in our simulations (*36*, *37*), then our estimates of the fold stabilization of cohesin by CTCF would change from approximately fourfold to approximately twofold, the CTCF stalling probability would change from 12.5 to 25%, and the looped state would now be held together by a single cohesin (fig. S16). We also note that TADs smaller than the 505-kb *Fbn2* TAD and TADs with stronger CTCF boundaries may have a higher looped fraction (*38*). Accordingly, we propose that our absolute quantification of the *Fbn2* looped fraction may now allow calibrated inference of absolute looped fractions genome wide on the basis of Micro-C (*13*).

Our findings reveal that the CTCF- and cohesin-mediated looped state that holds together CTCF boundaries of TADs is rare, dynamic, and transient. A key limitation of our study is that it represents just one loop in one cell type. Nevertheless, the *Fbn2* loop is among the strongest quartile of "corner peaks" in Micro-C maps, suggesting that most other similarly sized loops in mESCs are likely weaker than *Fbn2* (fig. S17). Our results thus rule out static models of TADs that exist in either a fully unlooped state or a fully looped state stably bridged by one cohesin (Fig. 1B). Instead, we show that the *Fbn2* TAD most often exists in a partially extruded state formed by a few cohesins in live cells (~92%; Fig. 4D), and that when the rare looped state is formed, it is transient (~10 to 30 min median lifetime; Fig. 4B). Because the partially extruded state dominates, this may be the functionally important TAD state. Thus, we suggest that CTCF-mediated transcriptional insulation may be mediated more by individual extrusion-blocking CTCF boundaries than by the rare fully looped state. Similarly, this suggests that regulatory interactions, such as those between enhancers and promoters, may depend more on frequent cohesin-mediated contacts within a TAD rather than rare CTCF-CTCF loops. This dynamic picture of TADs in live cells (Fig. 4D) may also help to explain cell-to-cell variation in 3D genome structure, and consequently stochasticity in downstream processes such as gene expression and cell differentiation.

REFERENCES AND NOTES

1. J. A. Beagan, J. E. Phillips-Cremins, *Nat. Genet.* **52**, 8–16 (2020).
2. M. Ganji *et al.*, *Science* **360**, 102–105 (2018).
3. Y. Kim, Z. Shi, H. Zhang, I. J. Finkelstein, H. Yu, *Science* **366**, 1345–1349 (2019).
4. I. F. Davidson *et al.*, *Science* **366**, 1338–1345 (2019).
5. S. Golfier, T. Quail, H. Kimura, J. Brugués, *eLife* **9**, e53885 (2020).
6. G. Fudenberg *et al.*, *Cell Rep.* **15**, 2038–2049 (2016).
7. O. Symmons *et al.*, *Genome Res.* **24**, 390–400 (2014).
8. M. V. Arrastia *et al.*, *Nat. Biotechnol.* **40**, 64–73 (2022).
9. Q. Szabo *et al.*, *Sci. Adv.* **4**, eaar8082 (2018).
10. T. J. Stevens *et al.*, *Nature* **544**, 59–64 (2017).
11. E. H. Finn *et al.*, *Cell* **176**, 1502–1515.e10 (2019).
12. B. Bintu *et al.*, *Science* **362**, eaau1783 (2018).
13. D. I. Cattoni *et al.*, *Nat. Commun.* **8**, 1753 (2017).
14. H. B. Brandão, M. Gabriele, A. S. Hansen, *Curr. Opin. Cell Biol.* **70**, 18–26 (2021).
15. H. Chen *et al.*, *Nat. Genet.* **50**, 1296–1303 (2018).
16. J. M. Alexander *et al.*, *eLife* **8**, e41769 (2019).
17. Z. Hensel, X. Weng, A. C. Lagda, J. Xiao, *PLOS Biol.* **11**, e1001591 (2013).
18. N. Khanna, Y. Zhang, J. S. Lucas, O. K. Dudko, C. Murre, *Nat. Commun.* **10**, 2771 (2019).
19. E. de Wit *et al.*, *Mol. Cell* **60**, 676–684 (2015).
20. T. Germier *et al.*, *Biophys. J.* **113**, 1383–1394 (2017).
21. T. S. Hsieh *et al.*, *Mol. Cell* **78**, 539–553.e8 (2020).
22. A. S. Hansen *et al.*, *Mol. Cell* **76**, 395–411.e13 (2019).
23. H. Hajjoul *et al.*, *Genome Res.* **23**, 1829–1838 (2013).
24. T. Natsume, T. Kiyomitsu, Y. Saga, M. T. Kanemaki, *Cell Rep.* **15**, 210–218 (2016).
25. S. S. P. Rao *et al.*, *Cell* **171**, 305–320.e24 (2017).
26. E. P. Nora *et al.*, *Cell* **169**, 930–944.e22 (2017).
27. G. Wutz *et al.*, *EMBO J.* **36**, 3573–3599 (2017).
28. T.-H. S. Hsieh *et al.*, Enhancer-promoter interactions and transcription are maintained upon acute loss of CTCF, cohesin, WAPL, and YY1. bioRxiv 2021.07.14.452365 [Preprint] (2021); https://www.biorxiv.org/content/10.1101/2021.07.14.452365v1.
29. A. Tedeschi *et al.*, *Nature* **501**, 564–568 (2013).
30. J. H. I. Haarhuis *et al.*, *Cell* **169**, 693–707.e14 (2017).
31. C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)* (Springer-Verlag, 2006).
32. Y. Li *et al.*, *Nature* **578**, 472–476 (2020).
33. G. Wutz *et al.*, *eLife* **9**, e52091 (2020).
34. C. Cattoglio *et al.*, *eLife* **8**, e40164 (2019).
35. A. S. Hansen, I. Pustova, C. Cattoglio, R. Tjian, X. Darzacq, *eLife* **6**, e25776 (2017).
36. E. Kim, J. Kerssemakers, I. A. Shaltiel, C. H. Haering, C. Dekker, *Nature* **579**, 438–442 (2020).
37. H. B. Brandão, Z. Ren, X. Karaboja, L. A. Mirny, X. Wang, *Nat. Struct. Mol. Biol.* **28**, 642–651 (2021).
38. P. Mach *et al.*, Live-cell imaging and physical modeling reveal control of chromosome folding dynamics by cohesin and CTCF. bioRxiv 2022.03.03.482826 [Preprint] (2022); https://www.biorxiv.org/content/10.1101/2022.03.03.482826v1.
39. H. B. Brandão, ConnectTheDots: A chromosome locus pair tracking library, Zenodo (2022); https://zenodo.org/record/6224612.
40. S. Grosse-Holz, SGrosse-Holz/tracklib: initial release, Zenodo (2022); https://zenodo.org/record/6144781.
41. M. Gabriele *et al.*, Trajectory data for: Dynamics of CTCF and cohesin mediated chromatin looping revealed by live-cell imaging, Zenodo (2021); https://doi.org/10.5281/zenodo.5770531.

ACKNOWLEDGMENTS

We thank R. Tjian and X. Darzacq for hosting early parts of this work; J. Alexander for providing the RNA destabilization elements; L. Joh for assistance with cloning; and members of the Hansen, Zechner, and Mirny laboratories for insightful discussions. **Funding:** This work was supported by the National Institutes of Health (grants R00GM130896, DP2GM140938, and R33CA257878 to A.S.H.; grant UM1HG011536 to A.S.H. and L.M.; and grant R01GM114190 to L.M.); the National Science Federation (grant 2036037 to A.S.H.); the Mathers' Foundation (A.S.H.); a Pew-Stewart Cancer Research Scholar grant (A.S.H.); Chaires d'excellence internationale Blaise Pascal (L.M.); an American-Italian Cancer Foundation research fellowship (M.G.); and core funding from the Max Planck Institute of Molecular Cell Biology and Genetics (C.Z.). **Author contributions:** A.S.H. conceived and initiated the project. H.B.B., M.G., S.G.H., L.M., C.Z., and A.S.H. designed the project. A.S.H. performed genome editing and generated the cell lines. G.M.D. cloned plasmids. M.G., A.J., C.C., and A.S.H. characterized and validated the cell lines. T.H.S.H. performed Micro-C. C.C. performed ChIP-Seq. M.G., A.J., and H.B.B. optimized imaging experiments with input from A.S.H. M.G., A.J., and H.B.B. collected the imaging data, with acquisition led by M.G. and A.J.; M.G. and A.J. performed control experiments and characterized the AID cell lines. H.B.B. developed the image-processing pipeline and analyzed the imaging data, with input from A.S.H., S.G.H., M.G., and A.J.; H.B.B. and S.G.H. developed the convolutional neural network, with input from A.S.H., M.G., and A.J.; H.B.B. performed polymer simulations with input from S.G.H. and L.M. M.G., A.J., H.B.B., and A.S.H. annotated trajectory data. S.G.H. and C.Z. designed BILD with input from H.B.B., L.M., and A.S.H. S.G.H. developed and benchmarked BILD, applied BILD to trajectory data, and developed MSD analysis with input from H.B.B., L.M., A.S.H., and C.Z. H.B.B. and S.G.H. analyzed polymer simulations. A.S.H., L.M., and C.Z. supervised the project. H.B.B., M.G., S.G.H., A.J., and A.S.H. drafted the manuscript and figures. All authors edited the manuscript and figures. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Cell lines, plasmids, and other materials are available upon request. All cell lines used in this study use the ANCHOR3 system for DNA locus visualization licensed from NeoVirTech. Recipients of the cell lines will need to execute a materials transfer agreement with NeoVirTech. The code and software used in this study, including the image-processing framework *ConnectTheDots* (*39*) and *BILD* (part of tracklib) (*40*), are freely available as described in table S5. The Micro-C and ChIP-Seq data associated with this study are available at the Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/) under accession number GSE187487. The raw trajectory data are available through Zenodo (*41*). For a list of all datasets used in this study, please see table S6.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abn6583
Materials and Methods
Supplementary Text
Figs. S1 to S17
Tables S1 to S6
References (*42*–*90*)
Movies S1 to S4
MDAR Reproducibility Checklist

12 December 2021; accepted 31 March 2022
10.1126/science.abn6583

HYDROGELS

Piezoionic mechanoreceptors: Force-induced current generation in hydrogels

Yuta Dobashi[1,2,3], Dickson Yao[1], Yael Petel[4], Tan Ngoc Nguyen[1,5], Mirza Saquib Sarwar[1,5], Yacine Thabet[1], Cliff L. W. Ng[1,4], Ettore Scabeni Glitz[1], Giao Tran Minh Nguyen[6], Cédric Plesse[6], Frédéric Vidal[6], Carl A. Michal[7,4], John D. W. Madden[1,2,5]*

The human somatosensory network relies on ionic currents to sense, transmit, and process tactile information. We investigate hydrogels that similarly transduce pressure into ionic currents, forming a piezoionic skin. As in rapid- and slow-adapting mechanoreceptors, piezoionic currents can vary widely in duration, from milliseconds to hundreds of seconds. These currents are shown to elicit direct neuromodulation and muscle excitation, suggesting a path toward bionic sensory interfaces. The signal magnitude and duration depend on cationic and anionic mobility differences. Patterned hydrogel films with gradients of fixed charge provide voltage offsets akin to cell potentials. The combined effects enable the creation of self-powered and ultrasoft piezoionic mechanoreceptors that generate a charge density four to six orders of magnitude higher than those of triboelectric and piezoelectric devices.

Human touch perception relies on soft cutaneous mechanoreceptors that intricately interface with even softer neural tissues. In advanced prosthetics, robotics, and wearable devices, the presence of sensor arrays that mimics human skin has led to elastomer-based piezoresistive (*1*), capacitive (*2*), and electret (*3*) sensors and electronic skins (*4*–*6*). These sensors enable soft interactions and can overcome the device-to-tissue mechanical mismatch that is inherent to conventional electronic devices. Ionic skin presents an alternative to electronic skin, producing ultrastretchable, tough, and transparent devices, including capacitive sensors, from hydrogels (*7*–*9*). In this study, we used hydrogels to directly produce ionic currents in response to pressure gradients—a piezoionic approach to sensing. This method complements other recent iontronic approaches such as the use of supercapacitive gates and highly conductive hydrogel–conducting polymer hybrids, enabling the large ionic charge exchange at the low voltages needed for neural interfaces (*10*, *11*). Piezoionics are shown to offer similarly high–charge density and low-voltage operation. Like their biological counterparts, they rely on ionic currents. Unlike mechanoreceptors, in which a transmembrane Donnan potential is released by mechanical gating (Fig. 1A), piezoionic sensors operate by pressure-driven ion flux (Fig. 1B). This occurs either by streaming a single polarity of ions or by one polarity of ions being preferentially carried over the counterions through the polymer matrix, creating a net charge imbalance.

Streaming and other pressure-driven voltage-generation methods have been documented in ion conductors such as Nafion (*12*), polypyrrole (*13*), cartilage (*14*), and hydrogels (*15*–*17*). Upon deformation of such piezoionic matrices, the open-circuit voltage gradient, ∇V_{sense}, is proportional to the applied pressure gradient, ∇P (Fig. 2A). In the simplest case, the relationship can be described as $\nabla V_{sense} = \alpha \nabla P$, where α is the piezoionic coefficient (see full discussion in the supplementary materials). α ranges from 0.01 to 100 nV/Pa, depending on the ionic species and matrix properties (table S2). It is equivalent to the ratio L_o/σ proposed by de Gennes *et al.* (*12*) to describe field generation in ionomeric polymers exposed to a pressure gradient, where L_o represents the coupling between the applied pressure and the induced ionic current, and σ is the ionic conductivity.

To explore the molecular origins of this effect and investigate applications in sensing, we designed an indentation experiment in which the open-circuit voltage or short-circuit current was measured between the deformed and undeformed portions of the gels (Fig. 2A and fig. S1). Indentation is analogous to the effect of a finger or other object pressing into skin. Two voltage responses to a very light step compression of 20 kPa are shown in Fig. 2B for acrylamide (AAm)–based charged copolymers that contain acrylic acid (AA) and proton counterions. This ultra-absorbent and biocompatible hydrogel, which is ubiquitous in commercial uses, generates several millivolts as a result of streaming potentials (*15*). Transient response time depends on polymer content, with the lower polymer content (15% by weight, black curve) producing a faster rise (1 s) and decay (15 s) than the higher polymer content version (30%, blue curve), which exhibits a slow rise (15 s) and little decay. The halving of the polymer content changes compressibility, leading to much larger relative stress relaxation (fig. S5A). This suggests that the faster response in the lower-content hydrogel is due to a greater permeability of the hydrogel matrix, increasing within-pore flow that carries protons and, in turn, producing a faster rise. The decay in response is also faster as the flow completes sooner.

A Poisson-Nernst-Planck model was combined with poroelastic mechanics (see supplementary materials) to help interpret the transient behavior (purple and red model-fit curves in Fig. 2B). The model indicates that the application of a pressure gradient creates an initial ionic current, driven by the internal convection of water and protons within the pores. This movement of charge creates concentration and voltage gradients that, in turn, lead to currents that oppose the convection-driven currents (Fig. 2, C and D; movies S8 to S10; and fig. S19). Peak voltage is reached when the back current, composed of diffusive and electrophoretic fluxes, balances the convective current. When electrolyte flow subsides, the transient separation of charge decays. The decay is much slower in the denser polymer (Fig. 2B), owing to the higher flow resistance and longer poroelastic time constant. This time constant is given by $\tau_p = \frac{(1-2\nu)\eta L^2}{2(1-\nu)G\kappa}$, where η is the fluid viscosity within the pores of the hydrogel, L is the indenter size or characteristic length scale of deformation, G is the shear modulus, ν is Poisson's ratio, and κ is the matrix permeability. G increases and κ decreases with increasing polymer content, with both being very strong functions of polymer volume fraction. Overall, the time constant of flow gets longer as polymer content increases. The simulation fits with the expectation that differences in matrix permeability and stiffness account for variations in rise and decay behaviors (model response in Fig. 2B, dashed lines). This understanding can be exploited to create a family of decaying or nearly steady responses of either positive or negative polarity, as simulated in fig. S25. The polarity is reversed by changing the sign of the mobile ion—for example, by replacing AA with [2-(methacryloyloxy)ethyl]trimethylammonium chloride (MAETAC) (fig. S4). A rich set of voltage responses is obtained as displacements and forces are increased (figs. S3, S4, and S7), with voltage amplitude increasing and transient response time changing. A fully comprehensive set of reasons for the force dependence of the decay is unclear and will require careful characterization of the changes in permeability and viscoelastic response as a

[1]Advanced Materials and Process Engineering Laboratory, University of British Columbia, Vancouver, BC, Canada. [2]School of Biomedical Engineering, Faculty of Applied Science, University of British Columbia, Vancouver, BC, Canada. [3]Institute of Medical Science, University of Toronto, Toronto, ON, Canada. [4]Department of Chemistry, University of British Columbia, Vancouver, BC, Canada. [5]Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC, Canada. [6]CY Cergy Paris Université, CY Advanced Studies, LPPI, 5 mail Gay Lussac, Neuville sur Oise, F-95031 Cergy-Pontoise Cedex, France. [7]Department of Physics and Astronomy, University of British Columbia, Vancouver, BC, Canada.
*Corresponding author. Email: jmadden@ece.ubc.ca

function of deformation, as discussed in the supplementary materials.

The streaming potentials that generate the responses plotted in Fig. 2B are reliant on having a single mobile charge within the electrolyte, balanced by a fixed charge of opposite sign on the polymer matrix. However, recent results (*18*–*22*) show that voltages are generated even in the absence of fixed charge. This effect is confirmed in an uncharged polyacrylamide (pAAm) hydrogel swollen with 0 to 1.5 M aqueous sodium chloride, whose peak sense voltage and ionic conductivity both monotonically increase as a function of concentration (Fig. 2E). No model has been developed to explain this response. The most straightforward conclusion is that the chloride ions displace faster than the sodium ions under indentation. This could be due to several effects: filtration resulting from size differences between ions (*23*), differential affinity to the matrix, or fluidic hindrance effects in which ion size is comparable to boundary layer dimension (*24*). Given that pore sizes in the hydrogels are within a factor of 10 of solvated ion sizes, and that filtration of potassium and chloride is evident even in pAAm hydrogels with low polymer content (*23*), we hypothesized that the physical barrier to diffusion provided by the polymer matrix is a factor that influences the generation of net current.

To observe the relationship between ionic mobilities and piezoionic response, experiments were conducted in poly(vinylidene fluoride-*co*-hexafluoropropylene) (PVDF-HFP) solid polymer electrolyte swollen with 0.1 to 3.0 M lithium bis(trifluoromethanesulfonyl) imide (LiTFSI) in propylene carbonate (PC). The salt was chosen because the lithium and fluorine enable diffusion coefficients to be readily extracted by nuclear magnetic resonance (NMR) (see supplementary materials). As salt concentration increases, there is also a reduction in the solvation shell and hydrodynamic radius of the lithium ion, from a coordination number of four in its first shell at dilute concentrations down to two PC molecules at high concentrations (*25*). This reduction is apparent in the increase in lithium diffusion coefficient with concentration, and relative to $TFSI^-$, as shown in Fig. 2G. The reduction occurs despite increasing solution viscosity (fig. S17). The diffusion coefficients of the two ions cross over, with $TFSI^-$ dominant at low concentrations and Li^+ at high concentrations. There is a corresponding polarity shift in the piezoionic voltage. The amplitude and sign of the sensor response to a sinusoidal indentation (0.1 Hz, 5% compressive strain) are shown in Fig. 2H (based on data plotted in fig. S13, A to H). The voltage is positive at low concentrations but crosses over to negative polarity between 1.5 and 2.0 M—the same range in which diffusion coefficients cross over. The apparent match, within experimental uncertainty, between the diffusion crossover point and the zero-sensing concentration is notable and supports the theory that the more-mobile and less-hindered ions determine the sign of the voltage response.

This differential effect is not explained by Darcy's continuum flow theory, which assumes that all molecules are carried at the same speed as the solvent. Given that the flow within the hydrogels is in a transition regime, it should not be surprising that interaction with the backbone leads to differential ion transport rates and currents. We propose a mechanistic picture of piezoionic electromechanical response in hydrogels with two mobile ions, where the cations and anions do not move at the same speed as the solvent, leading to current and voltage generation. The mobile ions are subject to a hydrodynamic drag force from the convection of electrolyte within the pores. This is balanced by interactions with the stationary polymer that hinder transport. If the effective mobilities of the ions within the matrix, as inferred by diffusion coefficients D_+ for cations and D_- for anions, are the same as they would be without the additional drag due to the polymer, as given by D_{o+} and D_{o-}, then we expect that no voltage would be generated, as both ions are carried at the speed of the fluid, v_f. Additionally, if cations and anions are equally hindered by the polymer, we expect that no sense voltage would be generated. These expectations lead to a predicted piezoionic coefficient that is proportional to the difference in the two diffusion coefficients, $\alpha = -\frac{eN\kappa}{\sigma\eta}\left[\frac{D_+}{D_{o+}} - \frac{D_-}{D_{o-}}\right]$ (see the supplementary materials for derivation). Here, e is the electronic charge (in coulombs), N is the concentration (per cubic meter), κ is the matrix permeability (in square meters), and η is the fluid viscosity (in pascal seconds). The relationship suggests that when the diffusion coefficients of the two ions are approximately equal, the voltage generation will be zero (Fig. 2, G and H). It also suggests that by maximizing difference in ion size (and thereby drag due to the matrix), a larger sense voltage will be obtained. The use of bulk solution values for D_{o+} and D_{o-} is a starting point but does not provide a quantitative match between experiment and theory because the polymer affects the interactions between ions

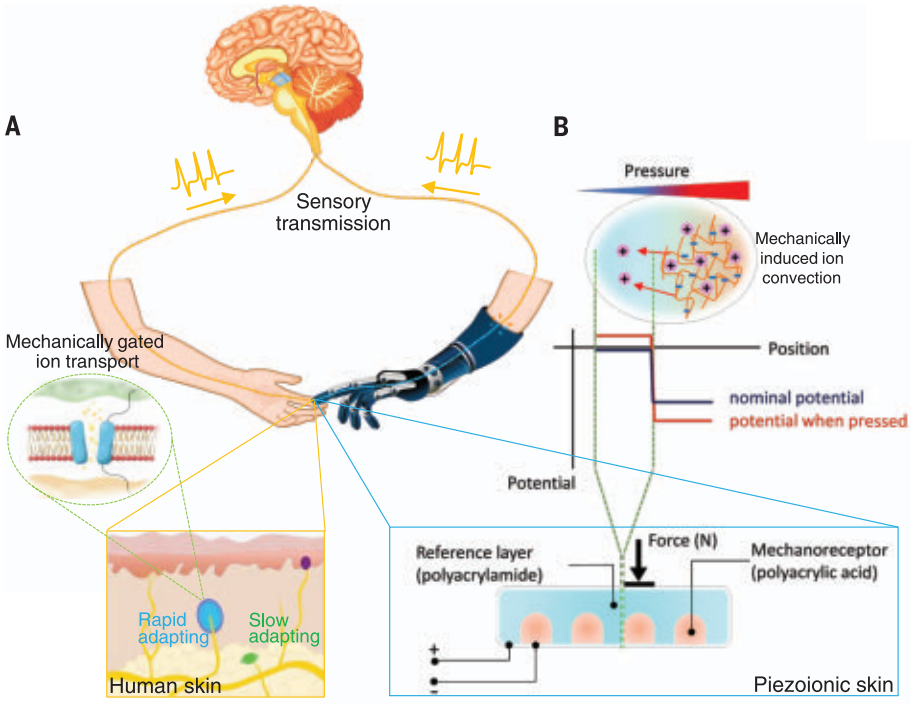


Fig. 1. Schematic representation comparing biological sensory transduction and piezoionics, and conceptual depiction of a piezoionic sensing device as an iontronic neuroprosthetic. (**A**) Mechanoreceptors have anchors that attach the membrane to the extracellular matrix as well as a cytoskeleton that both stretch an ion channel, causing an influx of sodium ions upon deformation. Once the cell membrane potential increases, voltage-sensitive ion channels (not shown) open to allow reversal to resting potential. The rapid- and slow-adapting mechanoreceptors differ, in part, in the mechanical structures of the cell. (**B**) Piezoionic skin assembly has a built-in potential difference set by the difference in fixed charge concentration between poly(acrylic acid) (polyAA; charged) and polyacrylamide (pAAM; neutral). Compression of the charged side creates a flux of solvent and protons that increases this potential difference.

and solvent. This is observed in the ionic conductivity, for which that of the polymer gel is higher than that of the bulk electrolyte at high salt concentrations (>1.0 M), peaking near the sensor crossover point (Fig. 2F), and the diffusion coefficients in bulk solution correspondingly drop below those in the matrix, crossing over at 1 M (fig. S14). The polymer matrix reduces ion-ion and ion-solvent interactions (*26*, *27*). The lower viscosity reduces drag, whereas effective ion concentration is increased by reducing ion pairing, hence increasing overall conductivity. This observation means that we cannot readily measure the hydrodynamic component of drag on ions within

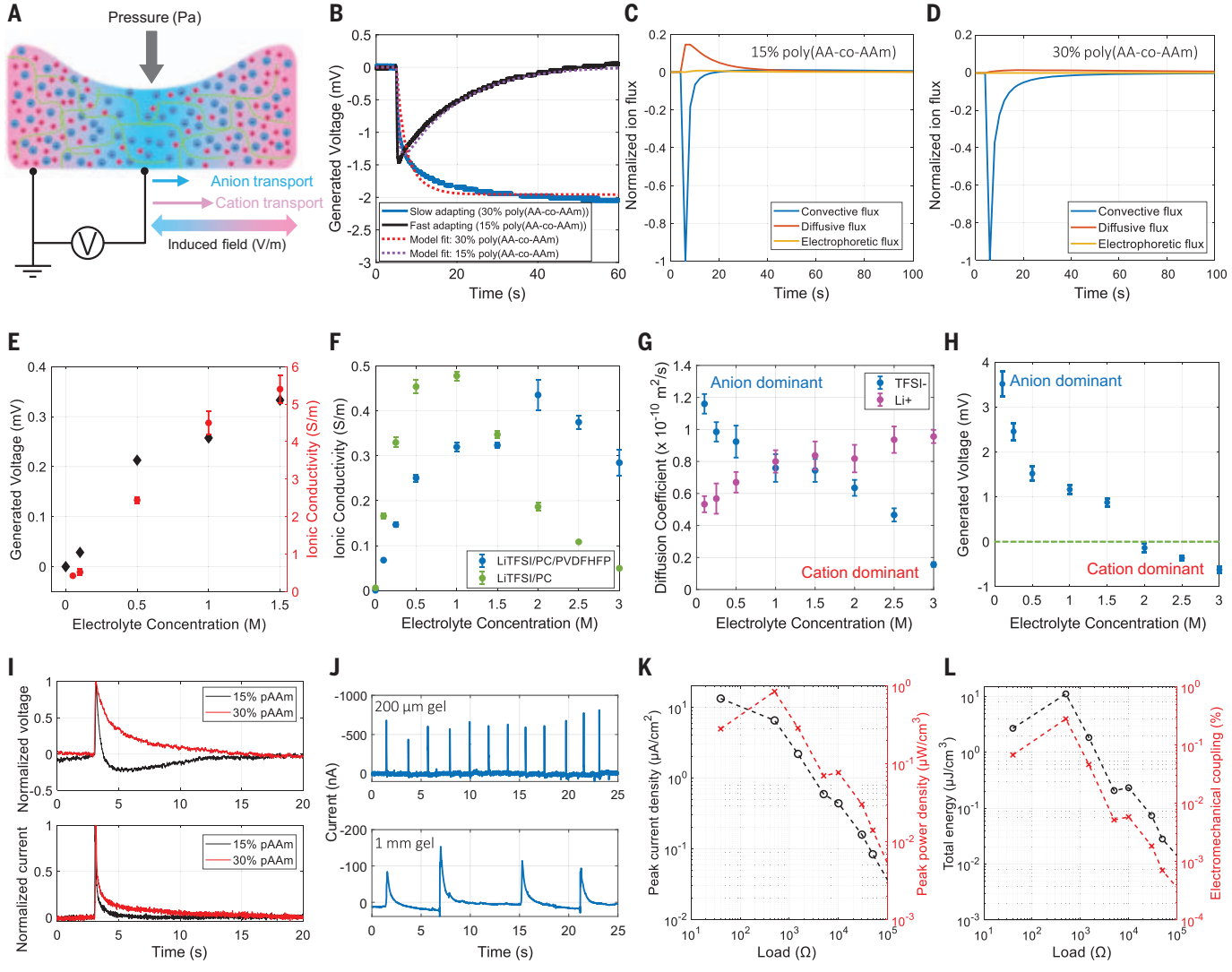


Fig. 2. Piezoionic mechanisms and electrical responses. (**A**) Schematic of a polymer gel under indentation, exhibiting differential ionic displacement and field, with smaller red cations being carried through the green polymer chain network faster than the blue anions, generating a charge imbalance and electric field. In the indentation experiments, a working (sense) electrode is placed under the indented portion, whereas a reference (ground) electrode is placed in the undeformed portion of the gel. In this case, a preferential displacement of the cations over anions results in a negative voltage reading. (**B**) Poly(acrylic acid-*co*-acrylamide) hydrogel voltage response upon step compression at 20 kPa. Transient behavior slows with increasing polymer content (15 and 30% w/v). Dashed fit curves were generated using a finite element simulation. (**C**) Simulation of the contributing components of ionic current at open circuit during the 15% poly(AA) gel compression, and (**D**) simulated breakdown of the ionic transport during the 30% poly(AA) gel compression, showing that convection dies off faster in the gel with lower polymer content. (**E**) Ionic conductivity of 15% w/v pAAm gel swollen with NaCl of varying concentrations and peak generated voltages (black; data are mean ± SD, N = 3) under 10% sinusoidal compression at 0.1 Hz, showing increased signal at higher concentrations (red; data are mean ± SD, N = 3). (**F**) Ionic conductivity of the PVDF-HFP solid polymer electrolytes (blue) and bulk LiTFSI/PC solution (green) as a function of concentration, showing that the polymer electrolyte is higher in conductivity at high concentrations (data are mean ± SD, N = 3). (**G**) Diffusion coefficients of Li^+ (pink) and $TFSI^-$ (blue) ions in the polymer electrolyte, measured by pulsed-field–gradient NMR, showing that the $TFSI^-$ is more diffusive at low concentrations, and Li^+ at higher concentrations (data are mean ± SD, N = 3). (**H**) Peak voltages generated by 0.1-Hz, 5% sinusoidal compressive strain in PVDF-HFP solid polymer electrolytes with PC/LiTFSI at varying concentrations, showing a reversal in sense voltage at a concentration similar to that of the diffusion crossover point (data are mean ± SD, N = 3). (**I**) Piezoionic transient response under a step compression from 15 and 30% pAAm swollen in 1.5 M NaCl. Note the faster decaying response both at open circuit and short circuit when the polymer content is lower. (**J**) Piezoionic transient response under repeated step compressions onto a 200-μm-thick (top) and a 1-mm-thick (bottom) pAAm gel swollen with 1.5 M NaCl, exhibiting decay times of ~50 ms and ~5 s, respectively, emulating rapid- and slow-adapting mechanoreceptors. (**K**) Peak current and power transfer as a function of resistive load, using a pAAm gel swollen with 1.5 M NaCl (data are mean ± SD, N = 3). (**L**) Energy harvested and electromechanical coupling as a function of external load, using a pAAm gel swollen with 1.5 M NaCl (data are mean ± SD, N = 3).

the pores. Nonetheless, the derived expression for piezoionic coefficient provides useful predictions regarding the effects of permeability, viscosity, ion size, concentration, and conductivity. For example, increased ion concentration leads to larger sense voltage (Fig. 2E).

A faster transient piezoionic response is observed in the pAAm-based hydrogel when polymer content is reduced (Fig. 2I). The higher fluid permeability (fig. S5) again leads to a faster poroelastic time constant in the 15% film than in the 30% film, as expected. A faster response is also observed as the gel thickness decreases. The current response time decreases from ~5 s at 1-mm thickness to 30 ms at 0.2 mm (Fig. 2J; close-up view of current response in fig. S10), owing to the local pressure spreading over a larger area as the hydrogel thickness increases, and hence the pressure gradient decreasing, as supported by simulation results in figs. S11 and S12, as well as movies S11 to S13. The transient response time varies greatly, from 30 ms at 0.2-mm thickness to >10 s at 2-mm thickness (fig. S9), similar to the response range (3 ms to 2 s) for rapidly- to ultraslowly-adapting human mechanoreceptors (*28*). Cutaneous mechanoreceptors attain different frequency responses by using specialized encapsulations around nerve endings to spatiotemporally modulate the

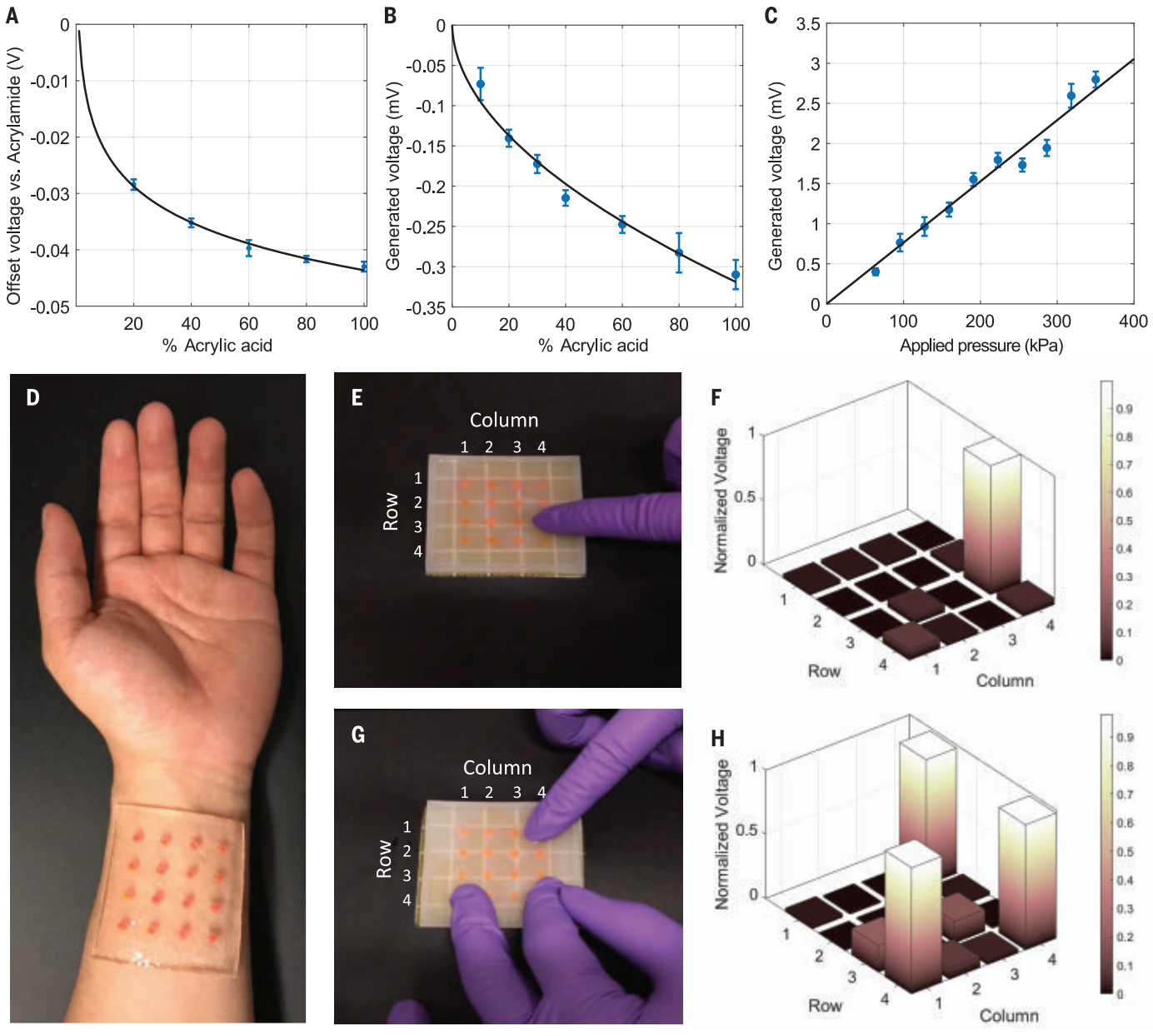


Fig. 3. Properties and operation of a 16-element piezoionic mechanoreceptor array. (**A**) Mechanoreceptor unit offset voltage as a function of the AA content, showing an increase in "resting" potential magnitude as backbone charge increases (data are mean ± SD, *N* = 3). (**B**) Indentation induced peak piezoionic voltages as a function of the AA content, showing that increased AA content results in increased magnitude of the generated voltage (data are mean ± SD, *N* = 3). (**C**) Peak generated voltage as a function of applied pressure (data are mean ± SD, *N* = 3). Fit lines represent smoothing spline fits for (A) to (C). (**D**) Photograph of a 16-element piezoionic mechanoreceptor array on a wrist, demonstrating conformability. (**E**) Photograph and (**F**) corresponding normalized voltage bar plot of the piezoionic mechanoreceptor array detecting a single touch (movie S1). (**G**) Photograph and (**H**) corresponding normalized voltage bar plot of the piezoionic mechanoreceptor array detecting multiple touches. (movie S2).

receptor deformation upon a given perturbation (*28*, *29*), whereas piezoionic current spikes are modified by permeability, modulus, and dimensions (as described by the poroelastic time constant; eq. S15).

Energy output of the piezoionic sensors is maximized by using classical impedance matching. Peak power transfer occurs when an external resistor of 500 ohms is employed (Fig. 2K), which matches the internal ionic resistance of the sensor. Perhaps unsurprisingly, these materials are not efficient energy harvesters. Electromechanical coupling was highest at 0.3% (Fig. 2L), as measured by applying a 10% compression at 1 Hz on a 1-mm-thick pAAm hydrogel swollen with 1.5 M NaCl. In comparison with stretchable piezoelectric and triboelectric devices (*3*, *9*), piezoionic voltage output is much lower, but peak power per volume is similar at 0.85 μW/cm^3 (Fig. 2L). Further exploration of dimensions, as well as electrolyte and polymer content, creates opportunities for investigating the upper limits of power density in piezoionic materials. These should scale well as dimensions are reduced, because the poroelastic time constant (eq. S15) is proportional to the square of the dimensions. Unlike piezoelectric (*30*) and triboelectric (*9*) generators, piezoionics can yield very high charges—in this case, 80 μC/cm^2 or 800 μC/cm^3. In PVDF, the charge per area is only 20 pC/cm^2 or 200 pC/cm^3 for the same sample dimensions and applied pressure. The highly stretchable electrets and triboelectric generators produce 4 nC/cm^2 and 0.01% of piezoionics in terms of charge per volume (*9*, *30*). Further comparisons of charge and power output are made with piezoelectric and triboelectric devices in the supplementary materials. The high charge density of piezoionics is particularly attractive for use in neural interfaces, where charge injections at tens of microcoulombs per square centimeter are routinely required while maintenance of low voltages is a key safety requirement (*31*).

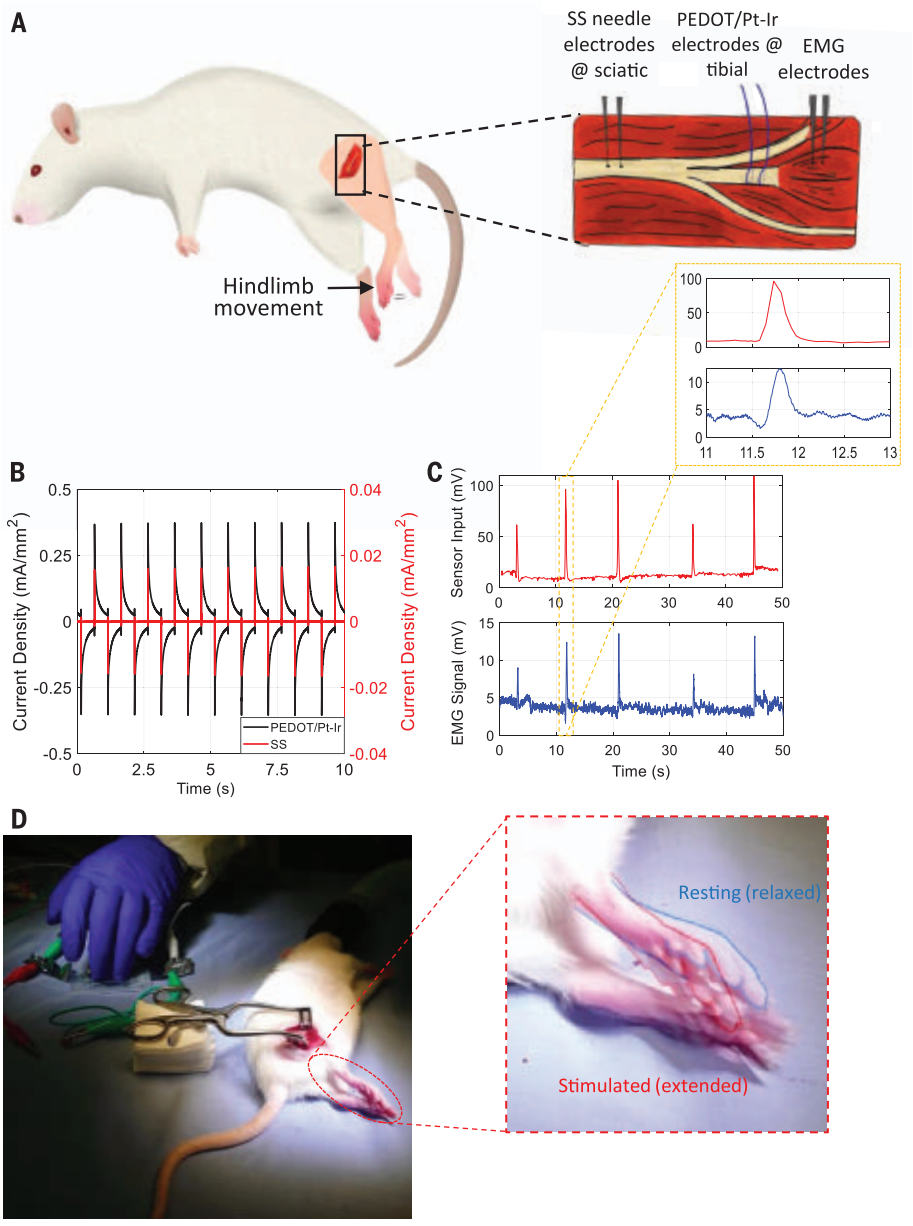


Fig. 4. Demonstration of piezoionic neuromodulation in a rodent model. (**A**) Setup for the piezoionic peripheral nerve stimulation experiment, showing electrode placement. (**B**) Area-normalized current responses to a square wave (peak-to-peak voltage = 5 mV) applied to the stainless-steel electrodes (red) and PEDOT-coated Pt-Ir electrodes (black) in a saline bath. Note the significant increase in the peak current as well as the steady-state transients using the PEDOT electrodes. (**C**) Sensor voltage and EMG signal reading during piezoionic stimulation, with stainless-steel electrodes at the sciatic nerve. The EMG signal lags the sensor input by around 5 to 8 ms as a result of propagation delay. (**D**) Piezoionic tibial nerve stimulation. Currents are generated by tapping on the piezoionic sensor directly, eliciting hindlimb movement (movie S5).

Piezoionics' similarities to mechanoreceptors in terms of charge, voltage, and time response led us to explore the creation of a piezoionic skin that can interact with the nervous system. We began by emulating transmembrane potential. Sodium-potassium pumps maintain a resting transmembrane potential of approximately −70 mV, which is partially depolarized and repolarized in a mechanoreceptor cell. The threshold potential change to produce an action potential is ~15 mV. To create a resting potential analog in the hydrogels, a boundary between a region of fixed polymer backbone charge and one of uncharged matrix was built into a hydrogel (Fig. 1B) The resulting artificial mechanoreceptor is composed of a 2-mm diameter, 2-mm-tall poly(AA-*co*-AAm) hemisphere that is surrounded in plane by pAAm swollen with 0.1 M KCl and formed into an array of 16 individually addressable touch sensors on an 8-mm pitch (Fig. 3D). An orange dye is used to identify the acrylic AA regions in the transparent sensor. In this case, protons are the only mobile ions, released by the dissociation of the AA. The immobility of the negative charge on the AA yields a negative Donnan potential relative to the AAm region, as protons diffuse into the pAAm plane. By increasing the ratio of AA to AAm units in the charged region, this built-in "resting potential" becomes as large as −50 mV (Fig. 3A). By applying a light pressure (0.5-Hz compressive sinusoidal indentation at 20% strain, corresponding to ~20 kPa) over the hemispheres, displaced water entrains the protons, and the potential becomes more negative, with larger response for a higher-AA content

(Fig. 3B). The pressure sensitivity of a single mechanoreceptor unit is 8 μV/kPa ($\alpha = 8 \times 10^{-9}$ V/Pa), probed with pressures of up to 360 kPa using a 0.1-Hz sinusoidal perturbation (Fig. 3C). A 4-cm–by–4-cm sensor array was tested under single- and multitouch conditions (Fig. 3, E to H and movies S1 and S2), as well as in detection of a swiping gesture (movie S3). A gentle finger press (~100 gram-force) generates voltage changes of approximately −10 mV, which are superimposed on the −50-mV Donnan potential. Little crosstalk between mechanoreceptors was observed. By changing the film thickness or polymer content or using current response instead of voltage, the properties of this transient response—including the decay time, which is ~1 s with the current dimensions—can be made faster or slower. The desired time constant depends on the application. If the aim is to detect change but ignore steady forces, then a fast time response is desired—for example, creating a situation in which the skin "forgets" about a constant force, just as we might stop being aware of clothing or a persistent noise.

This piezoionic skin may be suitable for soft robotics, medical devices, or wearable applications, given its mechanical compliance and self-powered nature. We demonstrate piezoionic detection of joint flexion and extension (fig. S27 and movie S6), for which the compliant nature makes it unobtrusive. We also detect the touch of a finger on skin and on a toy (movie S7). To be effective in such dry applications over hours or longer, the sensor needs a stretchable encapsulation layer. Ecoflex 00-30 was employed for this purpose (see supplementary materials), enabling up to several days of operation, which is sufficient for many wearable applications. Figures S29 and S30 demonstrate 390 cycles, where some drift is observed, returning to normal after resting, without degradation in performance. Switching to styrene-isobutylene-styrene block copolymers with low rates of water-vapor transmission should allow for months or even years of operation without dehydration (*32*), as could the use of lithium salts or ionic liquids (*33*). Owing to the bioinert nature of many synthetic hydrogels, it is also conceivable to employ piezoionics in implantable devices. Other advantages are the simplicity of the device and its adaptable compliance. An arterial phantom (fig. S28) demonstrates the sensor detecting a pulsatile flow that simulates cerebrovascular hemodynamics (*34*).

We performed peripheral nerve stimulation to demonstrate the possibility of self-powered piezoionic neuromodulation. A piezoionic sensor element composed of a pAAm hydrogel swollen in 1.5 M NaCl was pressed by hand to generate tactile-induced current pulses. Connections to the sciatic or tibial nerves of rodent models were made via conventional stainless-steel needle electrodes (250-μm diameter) or poly(3,4-ethylenedioxythiophene) (PEDOT)–coated (~1 μm) Pt-Ir wires (75-μm diameter), as shown in Fig. 4A. The PEDOT-coated Pt-Ir wires were gently wrapped around the tibial nerve, whereas the stainless-steel needles were placed by piercing the perineurium of the sciatic nerve. In either case, the distance between the contacts was 2 to 4 mm. Because the charge capacity of metals is relatively small [~0.2 F/m^2 (*35*)], conducting polymer PEDOT was used to increase charge injection capacity (Fig. 4B), as discussed in the supplementary materials. An electromyogram (EMG) was concurrently recorded by inserting a pair of needle electrodes in the gastrocnemius muscle.

A series of presses was then applied to the piezoionic sensor connected to the peripheral nerves. Using the stainless-steel electrodes, we did not observe a visible twitch of the hindlimb but did note a clear EMG signal, indicating that muscle is activated (Fig. 4C) but insufficient to attain a visual confirmation. The EMG signal lagged the voltage generated by the sensor by ~8 ms, as expected given the conduction velocities. Nonetheless, amplification of the sensor signal via analog-to-digital conversion into biphasic pulse trains enabled observation of hindlimb twitches (movie S4). Use of the PEDOT-coated Pt-Ir electrodes with the nerve, with the resulting currents in excess of 10 μA and well under 100 mV, produced clearly visible hindlimb twitches without any signal conditioning or amplification (Fig. 4D and movie S5).

Piezoionics offer sensing solutions that are soft, self-powering, and biocompatible. In response to pressure, they can be tailored to produce a wide temporal range of transient signals and provide much higher charges at much lower voltages than piezo- and triboelectric devices. Consequently, piezoionics are a good match for neural interfacing applications. Piezoionic devices are another iontronic tool, part of a system that, like nature, also includes computation, actuation, and energy storage.

REFERENCES AND NOTES

1. S. C. B. Mannsfeld *et al*., *Nat. Mater.* **9**, 859–864 (2010).
2. M. S. Sarwar *et al*., *Sci. Adv.* **3**, e1602200 (2017).
3. S. Zhang *et al*., *Nano Lett.* **20**, 4580–4587 (2020).
4. A. Chortos, J. Liu, Z. Bao, *Nat. Mater.* **15**, 937–950 (2016).
5. T. Someya, *IEEE Spectr.* **50**, 50–56 (2013).
6. D.-H. Kim *et al*., *Science* **333**, 838–843 (2011).
7. C.-C. Kim, H.-H. Lee, K. H. Oh, J.-Y. Sun, *Science* **353**, 682–687 (2016).
8. J.-Y. Sun, C. Keplinger, G. M. Whitesides, Z. Suo, *Adv. Mater.* **26**, 7608–7614 (2014).
9. X. Pu *et al*., *Sci. Adv.* **3**, e1700015 (2017).
10. W. Lee *et al*., *Adv. Mater.* **28**, 9722–9728 (2016).
11. Y. Liu *et al*., *Nat. Biomed. Eng.* **3**, 58–68 (2019).
12. P. G. de Gennes, K. Okumura, M. Shahinpoor, K. J. Kim, *Europhys. Lett.* **50**, 513–518 (2000).
13. W. Takashima, K. Hayasi, K. Kaneto, *Electrochem. Commun.* **9**, 2056–2061 (2007).
14. E. H. Frank, A. J. Grodzinsky, *J. Biomech.* **20**, 615–627 (1987).
15. A. Fiumefreddo, M. Utz, *Macromolecules* **43**, 5814–5819 (2010).
16. K. Prudnikova, M. Utz, *Macromolecules* **43**, 511–517 (2010).
17. K. Sawahata, J. P. Gong, Y. Osada, *Macromol. Rapid Commun.* **16**, 713–716 (1995).
18. Y. Liu *et al*., *Small* **12**, 5074–5080 (2016).
19. S. M. Villa *et al*., *ACS Macro Lett.* **8**, 414–420 (2019).
20. V. Woehling *et al*., *Multifunctional Materials* **2**, 045002 (2019).
21. M. S. us Sarwar *et al*., in *SPIE Smart Structures and Materials + Nondestructive Evaluation and Health Monitoring*, Y. Bar-Cohen, Ed. (International Society for Optics and Photonics, 2015), p. 943026; https://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=2239762.
22. Y. Dobashi, G. Allegretto, M. S. Sarwar, E. Cretu, J. D. W. Madden, *MRS Adv.* **1**, 63–68 (2016).
23. V. Kapur, J. Charkoudian, J. L. Anderson, *J. Membr. Sci.* **131**, 143–153 (1997).
24. K. B. Kosto, W. M. Deen, *Biophys. J.* **88**, 277–286 (2005).
25. H. Jiang *et al*., *Phys. Chem. Chem. Phys.* **21**, 10417–10422 (2019).
26. L. Othman, K. B. Md Isa, Z. Osman, R. Yahya, *Mater. Today Proc.* **4**, 5122–5129 (2017).
27. S. S. Sekhon, *Bull. Mater. Sci.* **26**, 321–328 (2003).
28. P. Delmas, J. Hao, L. Rodat-Despoix, *Nat. Rev. Neurosci.* **12**, 139–153 (2011).
29. M. Mendelson, W. R. Loewenstein, *Science* **144**, 554–555 (1964).
30. C. Dagdeviren *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **111**, 1927–1932 (2014).
31. D. B. McCreery, W. F. Agnew, T. G. H. Yuen, L. Bullara, *IEEE Trans. Biomed. Eng.* **37**, 996–1001 (1990).
32. S. Ebrahimi Takalloo *et al*., *Robotics* **8**, 60 (2019).
33. C. Keplinger *et al*., *Science* **341**, 984–987 (2013).
34. L. Zarrinkoob *et al*., *J. Cereb. Blood Flow Metab.* **35**, 648–654 (2015).
35. A. J. Bard, L. R. Faulkner, *Electrochemical Methods, Fundamentals and Applications* (Wiley, 1980).
36. Y. Dobashi *et al*., "Surface sensor arrays using ionically conducting material," US patent 10,401,241 (2017).
37. J. D. W. Madden *et al*., "Method and sensor for pressure sensing based on electrical signal generated by redistribution of mobile ions in piezo-ionic layer," US patent 10,663,358 (2018).

ACKNOWLEDGMENTS

This work greatly benefitted from discussions with V. X. D. Yang and J. Ku on the use of piezoionics in neural interfacing and monitoring. Peripheral neuromodulation studies were carried out with technical and methodological guidance from S. McErlane and R. Hildebrandt. We thank the reviewers for their insightful and constructive feedback. **Funding:** This study was supported by a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant and a Collaborative Research and Development Grant. J.D.W.M. was supported by the Institute of Advanced Studies at CY Cergy Paris Université through the Visiting Scholar program. **Author contributions:** Y.D. and J.D.W.M. conceived the experiments and concepts. Y.D. and D.Y. designed and fabricated the equipment for indentation experiments. Y.D., D.Y., T.N.N., M.S.S., Y.T., C.L.W.N., and E.S.G. synthesized the hydrogels and performed the indentation experiments. Y.D. and D.Y. processed and analyzed the data from the indentation experiments. Y.P. and C.A.M. constructed custom equipment and performed and analyzed the NMR measurements. J.D.W.M. and Y.D. interpreted the experimental data and conceived and developed the analytical and numerical models, with input from G.T.M.N., C.P., and F.V. Y.D. and D.Y. fabricated and characterized the sensor array. Y.D. and D.Y. fabricated the equipment and performed the rodent neuromodulation experiment. Y.D. and J.D.W.M. wrote the manuscript. All authors have critiqued and contributed to the revisions. **Competing interests:** Two US patents related to this work have been granted (*36*, *37*). **Data availability:** All data are available in the manuscript or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.aaw1974
Materials and Methods
Supplementary Text
Figs. S1 to S30
Tables S1 to S6
References (*38*–*58*)
Movies S1 to S13

15 March 2020; resubmitted 29 June 2021
Accepted 24 March 2022
10.1126/science.aaw1974

NEUROETHOLOGY

Food wanting is mediated by transient activation of dopaminergic signaling in the honey bee brain

Jingnan Huang[1]†, Zhaonan Zhang[1]†‡, Wangjiang Feng[1]†, Yuanhong Zhao[1]†, Anna Aldanondo[2], Maria Gabriela de Brito Sanchez[2], Marco Paoli[2], Angele Rolland[2], Zhiguo Li[1], Hongyi Nie[1], Yan Lin[1], Shaowu Zhang[3], Martin Giurfa[1,2,4]*§, Songkun Su[1]*§

The biological bases of wanting have been characterized in mammals, but whether an equivalent wanting system exists in insects remains unknown. In this study, we focused on honey bees, which perform intensive foraging activities to satisfy colony needs, and sought to determine whether foragers leave the hive driven by specific expectations about reward and whether they recollect these expectations during their waggle dances. We monitored foraging and dance behavior and simultaneously quantified and interfered with biogenic amine signaling in the bee brain. We show that a dopamine-dependent wanting system is activated transiently in the bee brain by increased appetite and individual recollection of profitable food sources, both en route to the goal and during waggle dances. Our results show that insects share with mammals common neural mechanisms for encoding wanting of stimuli with positive hedonic value.

The study of how brains encode pleasure (*1*, *2*) has uncovered the neural and molecular underpinnings of hedonic (i.e., pleasant) experiences such as reward. Psychological components in the processing of appetitive rewards have been identified (*3*) and traced to dissociable circuits in the mammalian brain (*1*): (i) "liking," which refers to the actual pleasurable impact of reward consumption; (ii) "wanting," which refers to the motivation to reach the reward; and (iii) "learning," which includes the implicit and explicit information about reward acquired through individual experience (*4*). In mammals, the process of wanting is mediated by neural systems that include mesolimbic dopamine (*1*, *5*).

Despite the fact that invertebrates exhibit sophisticated reward-seeking behaviors (*6*), the existence of an equivalent wanting system driving food search remains unknown. Honey bees are an attractive organism for the study of this topic owing to their intensive foraging activities, their developed social organization, and the presence of a symbolic code for information transfer about appetitive rewards (*7*) known as the waggle dance, which conveys information about direction and distance of profitable food sources (*8*). The existence of this code allows us to explore whether, besides communicating vectorial information pointing toward a food source, dancers transiently recollect the hedonic properties of the exploited food source. Does a wanting system drive foragers to profitable food sources and do waggle dances reactivate this system through the reminiscence of the reported food source? To answer these questions, we monitored foraging and dance behavior of identified foragers and simultaneously quantified and interfered with biogenic amine signaling in the bee brain during key phases of foraging and communication cycles. We demonstrate that dopamine levels increase in the brain of active dancing foragers both during dance initiation and upon departure toward an attractive food source. Individual hunger also elevates dopamine levels in the brain of foragers, thereby enhancing the hedonic value of sucrose solution and improving appetitive olfactory learning and memory retrieval. We thus provide evidence of the existence of a dopamine-based wanting system in the bee brain that is triggered by both individual hunger and personal recollection of the hedonic value of food sources in the context of the waggle dance.

We first trained honey bees (*Apis mellifera*) from an observation hive to collect sucrose solution from an artificial feeder. Trained, tagged bees flew regularly between the hive and the feeder for a period of 3 days. On day three, bees consistently reporting the presence of the feeder via regular waggle dances were identified. On day four, these bees were captured in different contexts and situations. At the feeder, tagged dancers were captured either upon arrival and before they started feeding or ~1 min after landing and starting feeding on the available sucrose solution. Within the hive, dancers were captured either immediately after dance initiation, upon dance ending (typically 50 to 60 s after dance start), or when departing for the food source (typically 10 to 20 s after dance end). We also collected "cessation" dancers, that is, bees that danced when the feeder was available but stopped dancing after the feeder was kept empty for 30 min.

Captured bees were immediately frozen in liquid nitrogen. High-performance liquid chromatography was performed on individual brains to quantify biogenic amine levels (nanograms per brain). Reliable measurements across all experimental conditions were obtained for dopamine (DA) and serotonin [5-hydroxytryptamine (5-HT)]. A highly significant interaction was found between the factors "biogenic amine" and "behavioral condition" (Fig. 1A) ($F_{4,130} = 15.53$, $P < 0.0001$). Whereas serotonin did not change with the phase of the foraging cycle or dance, significantly higher levels of dopamine were observed upon arrival at the food source before the bee started feeding and during waggle dance initiation (Tukey tests, $P < 0.001$). Although dopamine levels decreased at the end of dancing, foragers arrived at the feeder with increased levels of this amine. This increase was present upon departure from the hive, as shown by dopamine levels of foragers flying regularly between the hive and the feeder over the course of 3 days and captured on day four at the hive entrance after dance end and when preparing to leave the hive (Fig. 1B, "hive departure"). Equivalent levels of dopamine were found in these bees upon landing at the feeder (Fig. 1B, "feeder arrival"). Serotonin levels remained lower in both groups of bees. As a result, the interaction between the factors "biogenic amine" and "behavioral condition" was not significant ($F_{1,90} = 0.12$, $P = 0.73$). A significant variation was found for "biogenic amine" ($F_{1,90} = 261.32$, $P < 0.0001$) but not for "behavioral condition" ($F_{1,90} = 0.49$, $P = 0.49$). Overall, these results show that brain levels of dopamine increase when foragers start dancing in the hive to convey information about profitable food sources but then transiently decrease toward the end of the dance. Departure toward the reported food source is again accompanied by an increase of dopamine brain levels, whereas food ingestion at the goal decreases levels until the next dancing event. These fast, transient activations of the dopaminergic system could be indicative of an enhanced appetitive motivation upon food finding and recollection of food value through initiation of waggle dances.

The increased levels of dopamine in the brain of foragers may have been influenced by their 3-day cumulative flight experience prior to biogenic amine quantification. We therefore wondered whether the first experience with a profitable food source already enhanced dopamine levels. Marked bees were

[1]College of Animal Sciences (College of Bee Science), Fujian Agriculture and Forestry University, Fuzhou 350002, China. [2]Centre de Recherches sur la Cognition Animale, Centre de Biologie Intégrative (CBI), University of Toulouse, CNRS, UPS, 31062 Toulouse cedex 9, France. [3]Australian Research Council Centre of Excellence in Vision Science, Research School of Biology, College of Medicine, Biology and Environment, The Australian National University, Canberra, ACT 2601, Australia. [4]Institut Universitaire de France (IUF), Paris, France.
*Corresponding author. Email: martin.giurfa@univ-tlse3.fr (M.G.); susongkun@zju.edu.cn (S.S.)
†These authors contributed equally to this work.
‡Present address: Laboratoire Évolution et Diversité Biologique (EDB), UMR 5174, University of Toulouse, CNRS, UPS, 31062 Toulouse cedex 9, France.
§These authors contributed equally to this work.

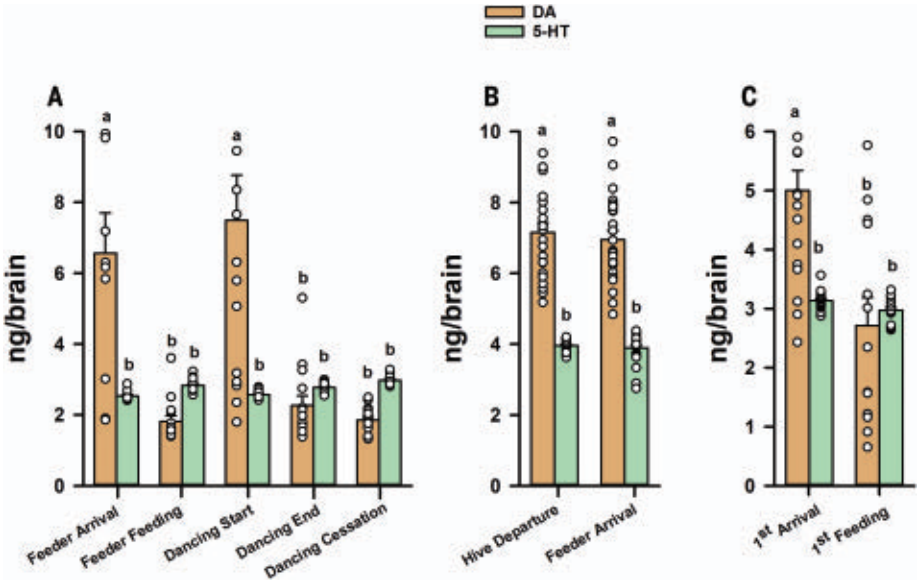


Fig. 1. Food expectation upon departure toward a food source and dance start increase dopamine levels in the honey bee brain. The three panels show dopamine (DA, orange bars) and serotonin (5-HT, green bars) levels (nanograms per brain; mean ± SEM) measured in individual brains of foragers belonging to different behavioral categories. Lowercase letters (a and b) above the bars indicate significant differences (Tukey tests; $P < 0.05$). Scatter plots show individual data distribution. (**A**) "Feeder arrival" corresponds to foragers that were observed dancing within the hive and that were captured upon arrival at the feeder reported in their dances and before they started feeding (DA: $n = 11$; 5-HT: $n = 10$). "Feeder feeding" corresponds to the same kind of bees captured 1 min after they landed on the feeder and started feeding (DA: $n = 15$; 5-HT: $n = 15$). "Dancing start" corresponds to bees that started dancing right after their return to the hive from the feeder and that were captured upon dance initiation (DA: $n = 15$; 5-HT: $n = 14$). "Dancing end" refers to the same kind of bees captured upon dance ending, ~1 min after dance start (DA: $n = 15$; 5-HT: $n = 15$). "Dancing cessation" refers to bees that stopped dancing after food was made unavailable for 30 min (DA: $n = 15$; 5-HT: $n = 15$). (**B**) "Hive departure" corresponds to foragers captured upon departure from the hive after dancing (DA: $n = 24$; 5-HT: $n = 24$). "Feeder arrival" corresponds to the same kind of bees captured upon arrival at the feeder and before they started feeding (DA: $n = 23$; 5-HT: $n = 23$). (**C**) "First arrival" shows foraging bees arriving at the feeder in their first active flight after prior appetitive experience at that location (DA: $n = 19$; 5-HT: $n = 19$). "First feeding" shows the same kind of bees captured 1 min after they landed on the feeder and started feeding (DA: $n = 13$; 5-HT: $n = 18$).

captured upon their first active visit (i.e., they flew on their own rather than being transported by the experimenter) to a feeder to which they were previously brought. Capture occurred either immediately upon arrival at the feeder or 1 min after landing and feeding. Figure 1C shows a significant interaction between the factors "biogenic amine" and "behavioral condition" ($F_{1,65} = 16.36$, $P < 0.001$). Whereas serotonin levels did not vary between the two experimental conditions (Tukey tests, not significant), dopamine levels were higher in bees arriving at the feeder for the first time ($P < 0.001$). Thus, specific expectations about profitable rewards are accompanied by high dopamine levels, irrespective of the amount of flight experience.

Having shown that waggle dances transiently enhance dopamine levels of dancers, we sought to determine whether the same effect was observable in dance followers that had never visited the feeder. We monitored waggle dances of marked foragers reporting the location of a trained feeder and captured nonmarked followers (i.e., bees with no feeder experience), either within 5 s after they started following the dancer ("short-term followers") or after 50 to 60 s ("long-term followers"). We also captured dancers in the middle of their dances (typically 20 to 30 s after dance start). We measured brain levels of biogenic amines and found that the interaction between the factors "biogenic amine" and "behavioral category" was not significant (fig. S1; $F_{2,100} = 0.17$, $P = 0.85$). Significant differences were found between dopamine and serotonin ($F_{1,100} = 41.26$, $P < 0.0001$) and between bee categories ($F_{2,100} = 16.38$, $P < 0.0001$). Dopamine levels were higher in dancers and increased in followers from the start to the end of the dance, yet this increase was not significant (Tukey tests). A similar trend was found for serotonin. Thus, dances might not activate per se a dopamine-based wanting system in followers.

To confirm the presence of a dopamine-based wanting system in foragers, we interfered with dopaminergic signaling and determined the effect of this interference on foraging. We trained marked foragers to collect sucrose solution from a feeder and measured their frequency of foraging bouts and their return time, that is, the length of time between two consecutive visits to the feeder, as these variables are sensitive indicators of appetitive motivation (*9*, *10*). Measurements were performed 30 min before topical delivery on the thorax (*11*) of the dopamine receptor antagonist fluphenazine (*12*, *13*), the dopamine receptor agonist 6,7-ADTN hydrobromide (hereafter, 6,7-ADTN) (*14*–*18*), or the solvent *N*,*N*′-dimethylformamide (DMF). Another control group remained untreated ("sham"). After 20 min, the same foraging variables were measured during a second 30-min period. The interaction between "treatment" and "measurement time" was not significant for the frequency of foraging bouts (Fig. 2A; analysis of variance for repeated measurements; $F_{3,82} = 1.57$, $P = 0.20$). However, a significant variation was observed between measurement times (before versus after: $F_{1,82} = 29.73$, $P < 0.0001$) and between treatments ($F_{3,82} = 3.40$, $P < 0.05$) as the antagonist fluphenazine decreased the frequency of bouts (Tukey test, $P < 0.01$). In the case of the return time, the interaction between "treatment" and "measurement time" was significant (Fig. 2B; $F_{3,82} = 7.49$, $P < 0.001$), and fluphenazine was again responsible for prolonging it ($P < 0.001$). The lack of effect of 6,7-ADTN on both variables measured may have been caused by a ceiling foraging motivation.

As fluphenazine might have affected motor activity (*19*) rather than foraging motivation, we repeated the previous experiment but decomposed the return time into (i) the time spent flying from the feeder to the hive, (ii) the time spent in the hive, and (iii) the time spent flying from the hive to the feeder. Foraging times were quantified in marked foragers before and after topical exposure to DMF, fluphenazine, or 6,7-ADTN. An untreated sham group was also included. Neither the time spent flying from the feeder to the hive (Fig. 2C; interaction between "time" and "group" factors: $F_{3,73} = 1.92$, $P = 0.13$) nor the time spent flying from the hive to the feeder (Fig. 2D; interaction between "time" and "group" factors: $F_{3,73} = 2.26$, $P = 0.09$) varied before and after treatment or between groups. On the contrary, the time spent within the hive exhibited a significant interaction (Fig. 2E; $F_{3,73} = 8.01$, $P < 0.001$). Specifically, fluphenazine increased the time spent in the hive (Tukey test, $P < 0.001$). Thus, blockade of dopaminergic signaling had no effect on flight activity but prolonged the time spent in the hive at the expense of foraging.

As honey bee foraging is determined by colony needs rather than by individual hunger (*20*), we asked whether individual appetitive

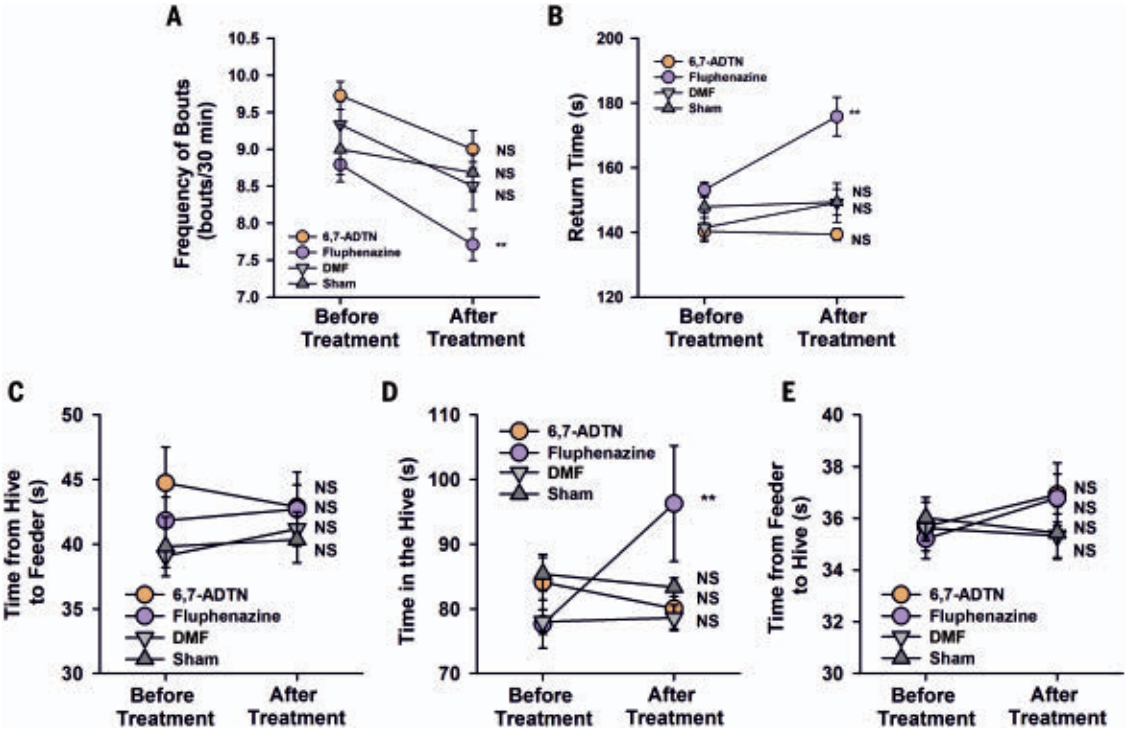


Fig. 2. Pharmacological blockade of dopaminergic signaling in honey bee foragers decreases the frequency of foraging and increases the time between successive visits to the feeder. Honey bees received a topical application on the thorax of either fluphenazine, 6,7-ADTN, or DMF or remained untreated (sham). The frequency of foraging bouts (**A**) and the time spent away from the feeder between two consecutive visits (return time) (**B**) were quantified during two 30-min periods, one before and one 20 min after topical exposure (6,7-ADTN: n = 22; fluphenazine: n = 24; DMF: n = 18; sham: n = 22). Treatment with fluphenazine significantly decreased the frequency of foraging bouts and increased the return time, thus revealing a negative impact on foraging motivation. (**C** to **E**) Pharmacological blockade of dopaminergic signaling in honey bee foragers increases the time spent in the hive at the expense of foraging. Neither the time spent flying from the feeder to the hive (C) nor from the hive to the feeder (D) was affected by the pharmacological treatments. On the contrary, blockade of DA signaling increased the time spent in the hive at the expense of the foraging time (E) (6,7-ADTN: n = 22; fluphenazine: n = 20; DMF: n = 22; sham: n = 13). In all panels, the values shown correspond to the mean ± SEM **P < 0.001; ns, not significant; Tukey tests.

needs also activate a dopamine wanting system. To this end, we studied whether starvation of individual foragers enhanced dopamine brain levels. Marked foraging bees of controlled age were captured at an artificial feeder and harnessed individually [see supplementary materials (SM) for details]. Bees were housed in an incubator for 1 or 2 hours and then frozen in liquid nitrogen for quantifying-dopamine and serotonin brain levels. The interaction between the factors "biogenic amine" and "starvation time" was significant (Fig. 3A; $F_{1,159}$ = 7.25, P < 0.01), as only dopamine (Tukey tests, P < 0.05), not serotonin, levels increased with starvation time. This increase is consistent with the activation of an individual wanting system.

We reasoned that manipulating dopamine levels would change the hedonic value of sucrose solution and thus individual appetitive responsiveness. To test this hypothesis, foragers captured upon landing at a feeder were individually enclosed within syringes, which allowed them to be fed with 30 μl of 1.0 M sucrose solution via a tip inserted in the syringe hub (*21*). In this way, we reduced individual appetitive responsiveness to be able to observe behavioral changes due to artificial dopamine increase. After feeding, bees were topically exposed to dopamine, 6,7-ADTN, or the solvent DMF. A sham group was also included. Bees were then individually harnessed and left at rest (see SM for details). They were then stimulated with six increasing concentrations of sucrose solution delivered to the antennae (0.1, 0.3, 1, 3, 10, and 30%, w/w) (*22*, *23*). The appetitive proboscis extension response to these stimuli was quantified, and an individual sucrose responsiveness score was calculated on the basis of the number of sucrose concentrations that elicited proboscis extension. Appetitive responsiveness varied between treatments (Fig. 3B; Kruskal-Wallis test; $H_{5,225}$ = 57.55, P < 0.0001). Whereas the two control groups (sham and DMF) showed similarly low responsiveness scores (mean rank comparison, not significant), the highest dose of both dopamine and 6,7-ADTN significantly increased sucrose scores (P < 0.0001). We confirmed this finding in bees availing themselves of an even higher energy supplement as they were individually fed with a mixture of honey, pollen, and sucrose solution (*24*–*26*) before the sucrose-responsiveness assay. Enhancing dopamine levels in these bees rescued their appetitive responsiveness (fig. S2; $H_{3,411}$ = 130.08, P < 0.0001), as their sucrose scores were significantly higher than those of the fed controls (P < 0.05).

We then tested whether manipulating dopamine would affect learning and retrieval of odorants predicting food. We took advantage of the olfactory conditioning of the proboscis extension response in which harnessed bees learn to associate a neutral odorant with a reward of sucrose solution (*27*, *28*). Foragers captured upon landing at a feeder were handled as in the previous experiments and topically exposed to DMF, the higher dose of dopamine, or the dopamine receptor antagonist flupentixol (*13*, *29*). An additional control was left untreated (sham). Each group was trained to discriminate 1-nonanol from 1-hexanol (*30*) (see SM for details).

All four groups learned the olfactory discrimination (fig. S3). To compare group performance, we calculated for each bee an acquisition score over the four last trials, which could vary between 2 and −2, with a higher positive score indicating successful learning. Scores varied

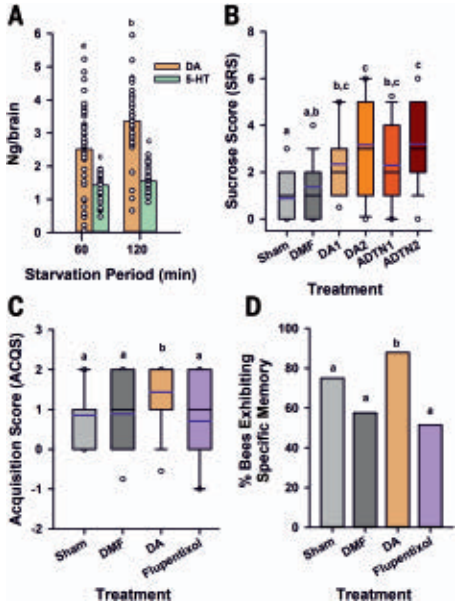


Fig. 3. An individual dopamine-dependent wanting system drives both appetitive responsiveness and appetitive learning and memory in honey bees. (**A**) Increased starvation enhances dopamine levels in the brain of foragers. Dopamine (DA) and serotonin (5-HT) measurements were performed in harnessed 23-day-old bees after immobilizing them and subjecting them to starvation periods of 1 or 2 hours (DA_{1h}: n = 42; DA_{2h}: n = 41; $5\text{-}HT_{1h}$: n = 39; $5\text{-}HT_{2h}$: n = 41). Increased hunger translates into higher DA levels in the bee brain. Lowercase letters (a, b, and c) above the bars indicate significant differences between groups (Tukey tests, $P < 0.05$). (**B**) Artificial increase of DA levels in honey bee foragers enhances their appetitive responsiveness to sucrose solution. Honey bee foragers captured at a feeder upon landing and before they started feeding were fed 30 µl of 1.0 M sucrose solution and then topically exposed on the thorax to one of two doses of either dopamine (DA_1: n = 29; DA_2: n = 30) or 6,7-ADTN ($ADTN_1$: n = 34; $ADTN_2$: n = 38). Control bees were treated with DMF (n = 47) or remained untreated (sham; n = 47). A sucrose responsiveness score (SRS) was computed for each bee subjected to an increasing series of six sucrose concentrations on the basis of its appetitive proboscis extension response (PER) in response to these stimulations. Higher SRSs indicate higher appetitive motivation. The highest dose of both DA and 6,7-ADTN significantly increased SRSs, thus revealing a higher appetitive motivation. For each treatment, the figure shows the median, 10th, 25th, 75th, and 90th percentiles as box plots with error bars; outliers are shown as 5th and 95th percentiles; the blue line within each box plot shows the mean. Lowercase letters (a and b) above the bars indicate significant differences between groups (Kruskal-Wallis test, post hoc mean rank comparisons, $P < 0.05$). (**C**) Artificial increase of DA levels in honey bee foragers enhances their appetitive olfactory learning. Honey bee foragers captured at a feeder upon landing and before they started feeding were harnessed and then topically exposed on the thorax to either dopamine (n = 48) or flupentixol (n = 54). Control bees were treated with DMF (n = 44) or remained untreated (sham; n = 48). Bees were trained along six trials to discriminate two odorants, one rewarded (CS+) and another not rewarded (CS−), using the olfactory conditioning of the PER. An acquisition score (ACQS) was computed for each bee on the basis of its responses to the conditioned odorants in the last two CS+ and last two CS− trials. Higher positive ACQSs indicate better learning performances. Bees in which DA was increased artificially demonstrated improved learning performance and exhibited significantly higher ACQSs. For each treatment, the figure shows the median, 10th, 25th, 75th, and 90th percentiles as box plots with error bars; outliers are shown as 5th and 95th percentiles; the blue line within each box plot shows the mean. Lowercase letters (a and b) above the bars indicate significant differences between groups (Kruskal-Wallis test, post hoc mean rank comparisons, $P < 0.05$). (**D**) Artificial increase of DA levels in honey bee foragers enhances their appetitive olfactory memory retrieval. Honey bee foragers captured at a feeder upon landing but before they started feeding were harnessed and trained to discriminate two odorants, one rewarded (CS+) and another not rewarded (CS−), using the olfactory conditioning of the PER. Bees that managed a successful discrimination in the last conditioning trial (i.e., responding to the CS+ and not to the CS−) were then topically exposed on the thorax to either dopamine (n = 33) or the flupentixol (n = 33). Control bees were treated with DMF (n = 33) or remained untreated (sham; n = 32). Memory retrieval was tested 1 hour after the last conditioning trial and 30 min after the topical exposure by presenting bees with the CS+ and the CS− in the absence of reward. The percentage of bees exhibiting specific memory (i.e., responding to the CS+ and not to the CS−) differed between treatments. It was significantly higher in the group treated with DA. Lowercase letters (a and b) above the bars indicate significant differences between groups (χ^2 contingency analysis, $P < 0.005$).

significantly between treatments (Fig. 3C; $H_{3,194}$ = 17.47, $P < 0.0001$). Bees with elevated dopamine levels learned better than the two control groups (Fig. 3C; $P < 0.05$) and the group treated with flupentixol (Fig. 3C; $P < 0.01$). Treatment with flupentixol tended to reduce discrimination at the end of training (fig. S3D), but the corresponding score did not differ from that of controls.

Finally, we determined whether manipulating dopamine brain levels would affect retrieval of an appetitive memory evoked 1 hour after the last conditioning trial. After conditioning foragers to discriminate the same previous two odorants, individuals that responded to the positive but not to the negative odorant in the last conditioning trial (fig. S4) were cold-anesthetized and assigned to different topical treatments (dopamine, flupentixol, or DMF). An additional group remained untreated (sham) (see SM for more details). Bees were then tested with both odorants, and the percentage of bees exhibiting specific memory [i.e., responding to the positive but not to the negative odorant (*30*)] was quantified for each group. This percentage varied significantly with the treatment (Fig. 3D; χ^2 = 12.47, df: 3, $P < 0.005$), as the group topically exposed to dopamine had a better retrieval performance than the other groups, which exhibited no differences in retrieval (χ^2 = 4.05, df: 2, not significant). Importantly, altering serotonin levels affected neither sucrose responsiveness (fig. S5A) nor olfactory appetitive learning (fig. S5B). Thus, the changes induced in appetitive responsiveness, learning, and memory are specific to dopaminergic signaling and point to a driving role for dopamine in an individual wanting system.

Our results suggest that distinct food-related motivational and communication states determine different levels of dopamine in the brain of honey bee foragers, consistent with the existence of a dopamine-based wanting system activated by both colony and individual appetitive needs. Foragers driven by an enhanced motivation for food exhibit higher dopamine levels. Accordingly, blockade of dopaminergic signaling results in a decrease of foraging activities, consistent with the inhibition of a wanting system mediated by dopaminergic signaling (*1*, *4*, *5*). Dopamine levels also increase when foragers report distance and direction of a profitable food source by means of the waggle dance and upon departure toward the appetitive goal. This finding suggests that dancers do not only report the vectorial information leading to the food source but also recollect transiently the appetitive properties of the food source at the start of dancing and when leaving the hive, thereby elevating dopamine levels in their brain. The waggle dance is thus more than an innate, automatic behavior (*7*), as it sets the framework for a dancer's transient evocation of food-source properties.

In addition to this socially embedded wanting system, individual hunger also leads to an increase in the level of dopamine in the forager brain. Accordingly, artificial elevation of dopamine increases the incentive value of sucrose solution and individual appetitive responsiveness and improves olfactory learning and memory retrieval. These results are consistent with alliesthesia modulation of incentive value by relevant physiological state (*31*) as, for instance, the improvement of memory retrieval upon dopamine elevation and in the absence of reward indicates that the incentive salience of a learned odor can be enhanced through an activation of the wanting system. Honey bees, unlike some mammals, are strictly eusocial

individuals with a sophisticated division of labor (*20*). As a result, their wanting system integrates both the social and the individual levels.

A neuroscience of pleasure has revealed that several common principles of the neural encoding of stimuli with positive hedonic value are shared among mammals (*1*). The presence of a dopamine wanting system in the honey bee brain suggests that a precursor to mammalian wanting systems may have evolved very early, that is, 220 million years ago, dating back to the oldest currently known Hymenoptera fossils (*32*).

REFERENCES AND NOTES

1. K. C. Berridge, M. L. Kringelbach, *Neuron* **86**, 646–664 (2015).
2. K. C. Berridge, M. L. Kringelbach, *Curr. Opin. Neurobiol.* **23**, 294–303 (2013).
3. K. C. Berridge, T. E. Robinson, *Trends Neurosci.* **26**, 507–513 (2003).
4. K. C. Berridge, T. E. Robinson, *Am. Psychol.* **71**, 670–679 (2016).
5. S. Robinson, S. M. Sandstrom, V. H. Denenberg, R. D. Palmiter, *Behav. Neurosci.* **119**, 5–15 (2005).
6. W. J. Bell, *Annu. Rev. Entomol.* **35**, 447–467 (1990).
7. K. von Frisch, *The Dance Language and Orientation of Bees*, L. Chadwick, Transl. (Harvard Univ. Press, 1967).
8. T. A. Sebeok, *Signs: An Introduction to Semiotics* (Univ. of Toronto Press, ed. 2, 2001).
9. J. A. Núñez, *J. Apic. Res.* **21**, 139–150 (1982).
10. J. A. Núñez, M. Giurfa, *Bee World* **77**, 182–196 (1996).
11. A. B. Barron, J. Maleszka, R. K. Vander Meer, G. E. Robinson, R. Maleszka, *J. Insect Physiol.* **53**, 187–194 (2007).
12. V. Vergoz, E. Roussel, J. C. Sandoz, M. Giurfa, *PLOS ONE* **2**, e288 (2007).
13. I. C. Kokay, A. R. Mercer, *Brain Res.* **706**, 47–56 (1996).
14. M. A. Humphries *et al.*, *J. Neurobiol.* **55**, 315–330 (2003).
15. R. Scheiner, S. Plückhahn, B. Oney, W. Blenau, J. Erber, *Behav. Brain Res.* **136**, 545–553 (2002).
16. S. E. Hewlett, J. D. Delahunt Smoleniec, D. M. Wareham, T. M. Pyne, A. B. Barron, *PLOS ONE* **13**, e0205686 (2018).
17. S. Akasaka, K. Sasaki, K. Harano, T. Nagao, *J. Insect Physiol.* **56**, 1160–1166 (2010).
18. J. A. Mustard *et al.*, *Brain Res. Mol. Brain Res.* **113**, 67–77 (2003).
19. J. A. Mustard, P. M. Pham, B. H. Smith, *J. Insect Physiol.* **56**, 422–430 (2010).
20. T. D. Seeley, *The Wisdom of the Hive: The Social Physiology of Honey Bee Colonies* (Harvard Univ. Press, 1995).
21. G. de Brito Sanchez *et al.*, *Sci. Rep.* **11**, 18311 (2021).
22. T. Pankiw, R. E. PageJr., *J. Comp. Physiol. A* **185**, 207–213 (1999).
23. R. Scheiner, R. E. Page, J. Erber, *Apidologie* **35**, 133–142 (2004).
24. L. Bestea *et al.*, *iScience* **25**, 103619 (2021).
25. L. Desmedt, L. Hotier, M. Giurfa, R. Velarde, M. G. de Brito Sanchez, *Sci. Rep.* **6**, 31809 (2016).
26. L. Bestea *et al.*, *Biol. Lett.* **18**, 20210520 (2022).
27. M. E. Bitterman, R. Menzel, A. Fietz, S. Schäfer, *J. Comp. Psychol.* **97**, 107–119 (1983).
28. M. Giurfa, J. C. Sandoz, *Learn. Mem.* **19**, 54–66 (2012).
29. W. Blenau, J. Erber, *Behav. Brain Res.* **96**, 115–124 (1998).
30. Y. Matsumoto, R. Menzel, J. C. Sandoz, M. Giurfa, *J. Neurosci. Methods* **211**, 159–167 (2012).
31. J. Zhang, K. C. Berridge, A. J. Tindell, K. S. Smith, J. W. Aldridge, *PLOS Comput. Biol.* **5**, e1000437 (2009).
32. R. S. Peters *et al.*, *Curr. Biol.* **27**, 1013–1018 (2017).
33. J. Huang *et al.*, Food wanting is mediated by transient activation of dopaminergic signaling in the honeybee brain, version 1, Figshare (2022); https://doi.org/10.6084/m9.figshare.17912468.v1.

ACKNOWLEDGMENTS

We thank three anonymous reviewers for useful comments and suggestions on a previous version of the manuscript. **Funding:** This research was supported by the National Natural Science Foundation of China (31772684) and China Agriculture Research System of MOF and MARA (CARS-44). Support was also provided by the CNRS, the University Paul Sabatier of Toulouse, the Institut Universitaire de France, and the Fujian Agriculture and Forestry University. **Author contributions:** J.H., Z.Z., W.F., and Y.Z. performed the experiments relating aminergic signaling with foraging and dance activity and also contributed to the analysis of these data, together with Z.L., H.N., Y.L., S.Z., and S.S., who supervised data collection. A.A., A.R., M.P., and M.G.d.B.S. performed the experiments relating individual aminergic signaling and individual appetitive responses and learning and also contributed to data analysis, together with M.G., who supervised data collection. S.S. and M.G. conceived of the project and supervised the research work. Resources and funding were provided by S.S. and M.G. M.G. wrote the original draft of the manuscript. All authors discussed the data and contributed to the revision and editing of the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Data are available in Figshare (*33*).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abn9920
Materials and Methods
Figs. S1 to S5
References (*34*–*44*)
MDAR Reproducibility Checklist

6 January 2022; accepted 4 March 2022
10.1126/science.abn9920

EPIDEMIOLOGY

Rabies shows how scale of transmission can enable acute infections to persist at low prevalence

Rebecca Mancy[1], Malavika Rajeev[2], Ahmed Lugelo[3,4], Kirstyn Brunker[1], Sarah Cleaveland[1], Elaine A. Ferguson[1], Karen Hotopp[1], Rudovick Kazwala[3], Matthias Magoto[5], Kristyna Rysava[6], Daniel T. Haydon[1], Katie Hampson[1]*

How acute pathogens persist and what curtails their epidemic growth in the absence of acquired immunity remains unknown. Canine rabies is a fatal zoonosis that circulates endemically at low prevalence among domestic dogs in low- and middle-income countries. We traced rabies transmission in a population of 50,000 dogs in Tanzania from 2002 to 2016 and applied individual-based models to these spatially resolved data to investigate the mechanisms modulating transmission and the scale over which they operate. Although rabies prevalence never exceeded 0.15%, the best-fitting models demonstrated appreciable depletion of susceptible animals that occurred at local scales because of clusters of deaths and dogs already incubating infection. Individual variation in rabid dog behavior facilitated virus dispersal and cocirculation of virus lineages, enabling metapopulation persistence. These mechanisms have important implications for prediction and control of pathogens that circulate in spatially structured populations.

Understanding the processes that regulate endemic disease dynamics remains a long-standing challenge in epidemiology (*1*, *2*), and the mechanisms that enable long-term persistence at low prevalence remain largely unexplored (*3*). This is particularly true for canine rabies, a fatal zoonotic virus for which naturally acquired immunity has not been demonstrated. The basic reproductive number R_0 of rabies, which is defined as the expected number of secondary cases produced by a typical infectious individual in a fully susceptible population (*4*), is low (between 1.1 and 2) and is relatively insensitive to dog density (*5*), making the disease amenable to elimination through dog vaccination (*6*). Yet dog-mediated rabies remains endemic across Africa and Asia, where it kills tens of thousands of people every year (*7*), and its persistent circulation at such low prevalence in largely unvaccinated populations is an enduring enigma.

Rabies is, however, a usually tractable system for understanding how population-level patterns of infection emerge from pathogen transmission at the individual level. Rabies is transmitted through bites, which can often be observed, and the clinical signs are readily identifiable, with infected animals typically dying within 1 week of disease onset (Fig. 1E). Capitalizing on these distinctive characteristics, we conducted exhaustive contact tracing to generate spatially resolved data on rabies infection and transmission in Serengeti district, Northern Tanzania, between January 2002 and December 2015. Serengeti district adjoins other populated districts to the north and west and Serengeti National Park to the southeast (Fig. 1A). We previously showed that domestic dogs maintain rabies in this part of Tanzania, with infrequent spillover to wildlife and spillback to domestic dogs (*8*). In Serengeti district's population of around 50,000 dogs (and 250,000 people), we traced 3612 rabies infections (comprising 3081 cases in dogs, 75 in cats, 145 in wildlife, and 311 in livestock) (Fig. 1), along with 6684 potential transmission events to other animals and 1462 people bitten by rabid animals, of whom 44 died from rabies. Most identified cases could be statistically

[1]Institute of Biodiversity, Animal Health and Comparative Medicine, University of Glasgow, Glasgow, UK. [2]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA. [3]Department of Veterinary Medicine and Public Health, Sokoine University of Agriculture, Morogoro, Tanzania. [4]Ifakara Health Institute, Dar es Salaam, Tanzania. [5]Serengeti District Veterinary Office, Mugumu, Tanzania. [6]The Zeeman Institute for Systems Biology and Infectious Disease Epidemiology Research, University of Warwick, Warwick, UK.
***Corresponding author. Email: katie.hampson@glasgow.ac.uk**

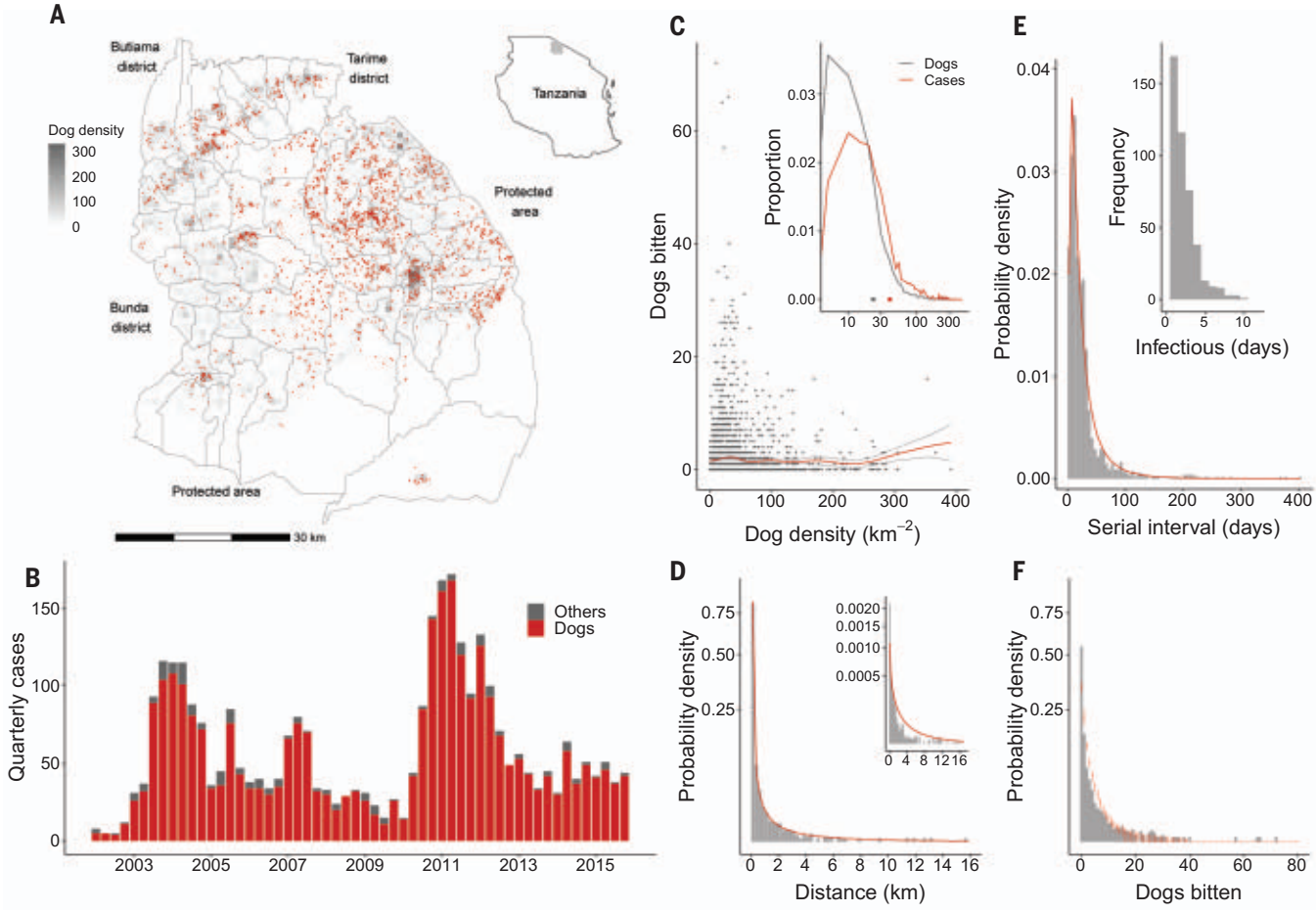


Fig. 1. Rabies in Serengeti district. (**A**) Mapped cases are in red. Shading indicates dog density, and lines indicate village boundaries. (Inset) The district location in Tanzania. (**B**) Monthly time series of cases in domestic dogs (red; n = 3081) and other carnivores (gray, n = 214). Species are detailed in (*9*). (**C**) Dogs bitten per rabid dog versus dog density at each rabid dog's location (1-km^2 scale), showing no apparent relationship. The red line indicates the generalized additive model prediction, and gray lines indicate the standard error; $P > 0.05$. (Inset) The proportional distribution of the dog population and of case locations in relation to dog density on a log scale. Squares indicate mean dog population density (black; 23 dogs/km^2) and mean dog density at case locations (red; 41 dogs/km^2), indicating higher per capita incidence in higher-density areas (independent samples t test on log-transformed data, $T = 22.45$, $P < 2.2 \times 10^{-16}$). (**D** and **E**) Distributions of (D) rabid dog step-lengths between contacts and (inset) distances to contacts and (E) serial intervals and (inset) infectious periods. The best-fitting distributions are in red (table S1), with step-lengths and distances censored for animals with unknown starting locations (*9*). (**F**) Dogs bitten per rabid dog from contact tracing (gray) and simulated from the individual-based model (red; 5th to 95th percentiles). The y axis is square root transformed in (D) and (F) to better illustrate extreme values.

linked to plausible progenitors, indicating that contact tracing detected most cases. Further analysis indicated that 83 to 95% of cases were detected, with missed cases mainly being those generating limited, if any, onward transmission (*9*). These data show that despite local vaccination effort (coverage varying between 10 and 40%) (fig. S3B), rabies circulated continuously, with a maximum prevalence of just 0.15%.

Endemic diseases are thought to be primarily regulated by the depletion of susceptible hosts, typically through disease-induced (or vaccine-acquired) immunity, counterbalanced by births of susceptibles and deaths (*4*). Yet the very low prevalence and absence of acquired immunity for rabies indicates that large-scale depletion of susceptible hosts is negligible, challenging this explanation as a mechanism for persistence. We estimated dog densities at high spatial resolution through a district-wide census, georeferencing almost 36,000 households and recording the vaccination status of dogs [in this setting, almost all dogs are owned but also free-roaming, and there are no feral dogs (*10*)]. Although we saw no clear relationship between dog population density and contact rate when examining dogs bitten per rabid dog (Fig. 1C), mapping rabies infections revealed a small yet significantly higher incidence of cases in higher-density areas (Fig. 1C), which suggests density-dependent processes. These observations are difficult to reconcile: How does transmission respond to dog population density, and what processes keep prevalence so low?

We propose that understanding the fine-scale structure of rabies transmission networks is critical to explaining its persistent dynamics and can inform its control and eventual elimination. We used the serial interval distribution—defined as the interval between the onset of infection in primary and secondary cases—and movement of traced rabid dogs to reconstruct transmission trees, comprising putative introductions into the district and descendent chains of transmission. Over the 14 years, we estimated around 238 introductions (8 to 24 per year) that led to onward circulation, most likely spreading from neighboring villages (movie S1). Twenty-two transmission chains, accounting for >70% of cases, circulated for more than 12 months (including two for more than 4 years), illustrating how co-circulation of lineages contributes to persistence within a metapopulation (Fig. 2). We applied an individual-based model, seeded

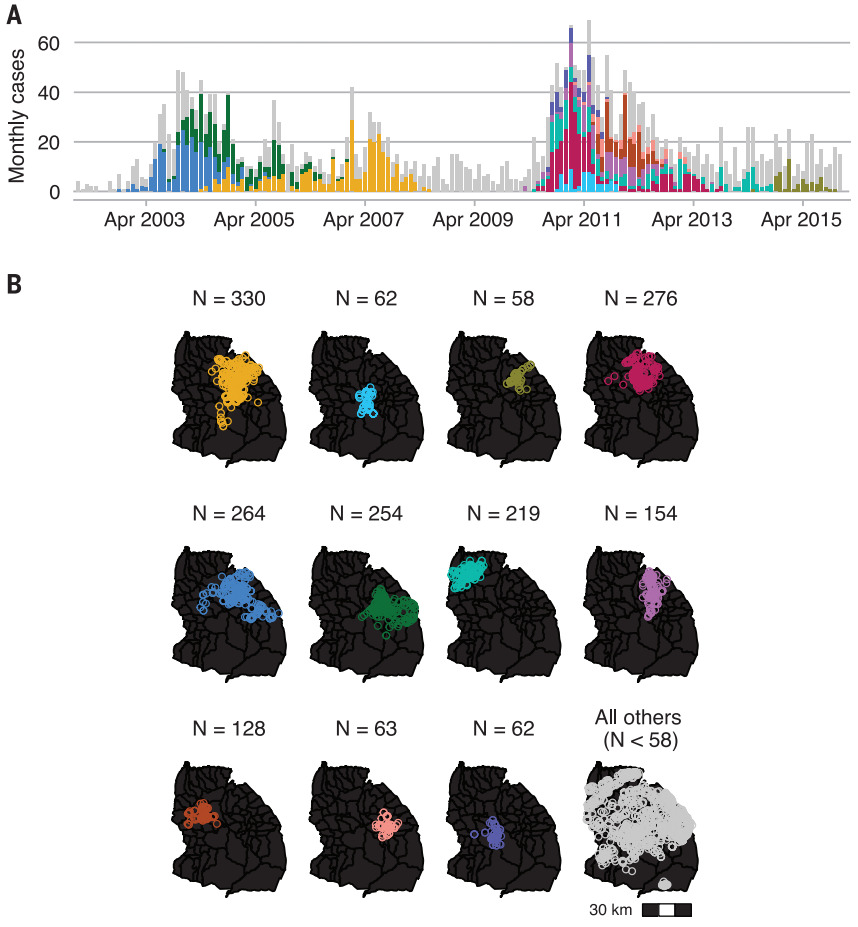


Fig. 2. The spatiotemporal distribution of transmission chains. (**A**) Monthly cases, estimated to be between 83 and 95% of all cases in the district. The 11 chains with most cases are highlighted in color, and all smaller chains (<58 cases) are shaded gray. (**B**) Spatial distribution of monthly cases, from the consensus tree of 1000 bootstrapped reconstructions (*9*).

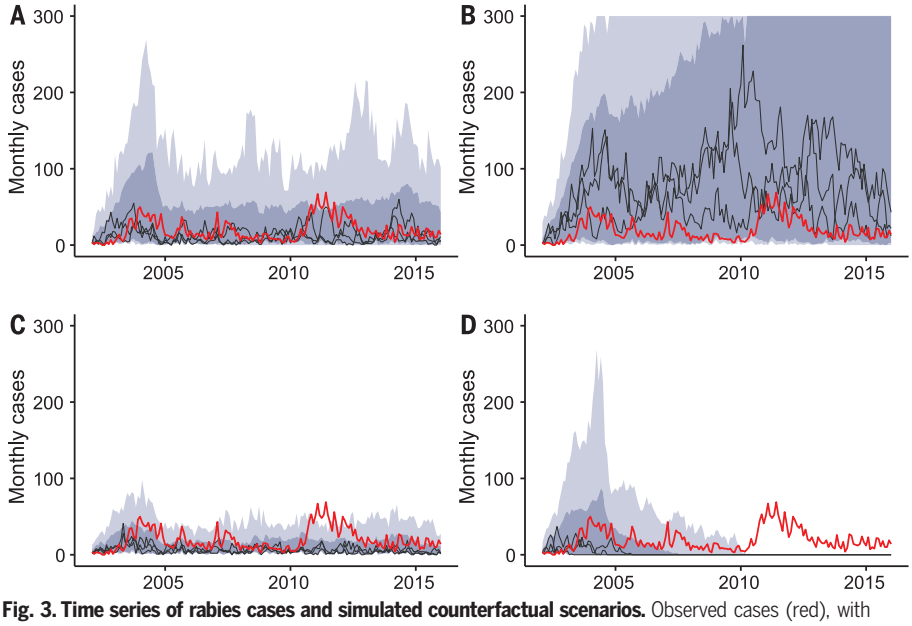


Fig. 3. Time series of rabies cases and simulated counterfactual scenarios. Observed cases (red), with interquartile (dark shading) and 95% (light shading) prediction intervals, both computed pointwise, from simulations at the 1-km^2 scale, and with three illustrative example runs (dark lines). The simulated scenarios are (**A**) with vaccination campaigns as implemented, (**B**) under low vaccination coverage, (**C**) with vaccination campaigns as implemented but no individual heterogeneity in contact parameters, and (**D**) without incursions after the first year.

by these introductions, to the spatially resolved case and dog density data to investigate the processes that modulate transmission, the scale over which they operate, and how they facilitate these metapopulation dynamics.

We examined the effect of host (dog) density on contact (biting) within our stochastic individual-based model, drawing from predator-prey functional response theory (*11*), to fit parameters defining the transmission process in relation to host density (*12*). We fitted the parameters assuming that susceptible and infected dogs were well mixed at different scales, ranging from the whole district to within 0.25-km^2 grid cells (fig. S2). The best-fitting parameters at each scale differed in their simulation outcomes. Only models at the 1-km^2 scale reliably reproduced observed dynamics and captured emergent population- and individual-level properties (Figs. 1F and 3A and fig. S10). We conclude that the processes that regulate the size of outbreaks and overall prevalence of rabies operate primarily on scales that are much smaller than those typically modeled for this disease.

Our modeling shows that epidemic growth is curtailed in two ways: from deaths of rabid dogs reducing contact opportunities and through redundant exposures of dogs already incubating infection. These processes increasingly stem transmission in lower-density areas, where dogs have fewer contact opportunities (Fig. 4). Simulations of index infections to estimate R_0 (in an entirely susceptible population) show that rabid dogs bite, on average, 2.91 dogs, leading to approximately 1.47 secondary cases per index case [95% percentile intervals (PI) 1.39 to 1.56]. This R_0 value is slightly higher than previously estimated (*5*), in part because of population growth (median dog density increased from 12 to >20 dogs/km^2 over the 14 years) but varies across the landscape in relation to dog densities (Fig. 4) and according to how it is measured; estimates simulated from dogs sampled from the landscape (by grid cell, rather than in proportion to density) were slightly lower (1.35, 95% PI 1.27 to 1.43), whereas estimates from dogs sampled from the transmission network were slightly higher (1.48, 95% PI 1.38 to 1.58) because more cases occur in higher-density grid cells (Fig. 4) (*9*).

Under endemic circulation, the effective reproduction number, R, declines by just over 30% and remains near 1. Around 78% of this reduction is from recent rabies deaths removing potential contacts, and 22% is from re-exposures of already incubating dogs. Thus, infected dog movement (fig. S1) determines the scale of mechanisms that regulate endemic dynamics, so that even small outbreaks (approximately five cases per square kilometer) can result in substantive reductions in R, given the heterogeneous distribution of

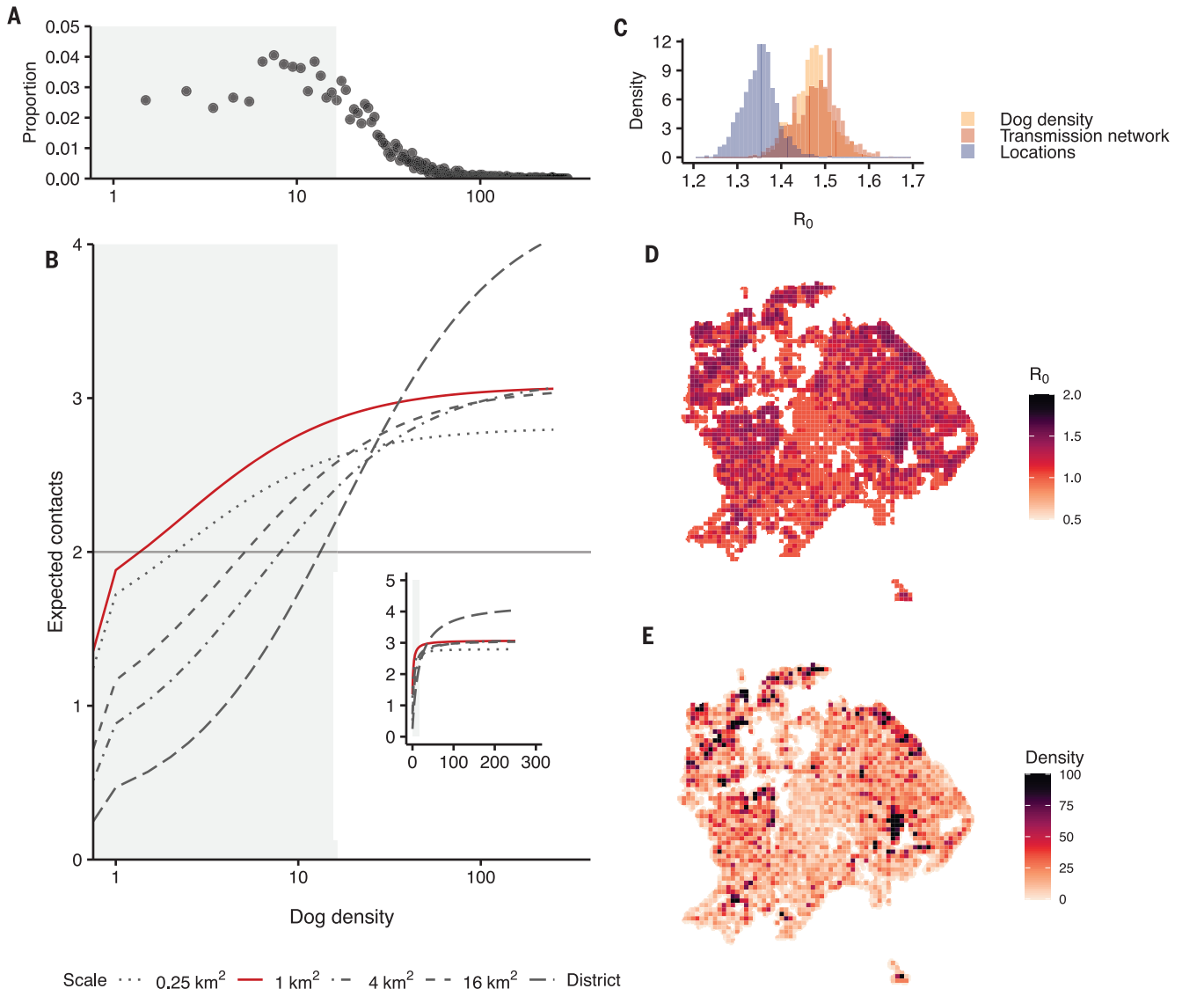


Fig. 4. Rabies transmission in relation to population density. (**A**) Distribution of dog densities in Serengeti district on a log scale. (**B**) Holling curves computed from the most reliable parameter set at each scale (*9*), with contact rate rescaled by the median infectious period (2 days) and dog density on a log scale. Gray shading indicates dog densities below the median (16.4 dogs per km^2) midway through the period (median density increased from 12 to 22 dogs/km^2 over the 14 years). (Inset) Replotted on a linear scale. At the optimal spatial scale (red line; 1 km^2), contact rates are density dependent at low dog densities and become increasingly density independent at higher densities. Parameter sets at other scales failed to reliably generate the observed dynamics. The Holling curve indicates how even removal of a small number of susceptible dogs locally (from rabies deaths or incubating infection) reduces transmission so that R approaches 1, which corresponds to around two contacts per infection (horizontal line), with half developing rabies. (**C**) Histogram of R_0 estimates from simulating index infections (*9*), either by location (blue), density (orange), or from the transmission network (red). (**D** and **E**) Mapped (D) R_0 estimates and (E) dog density at the midpoint.

dogs on the landscape (Fig. 4). There remains ambiguity in the mechanisms that underpin density-independent transmission at higher dog densities. Human responses likely play a role (45% of traced rabid dogs were either killed or tied) and are to some extent captured in our model, but these may operate differently during larger outbreaks (beyond those observed) and in more urbanized populations (<2% of dogs in this rural district live at densities >100 dogs/km^2). Our data highlight how better understanding of functional responses that describe theoretical relationships of transmission with density (*10*, *11*) are needed to predict endemic pathogen dynamics.

Individual variation in disease transmission causes rare but more explosive outbreaks and more frequent extinctions (*13*). We observed considerable variation in rabid dog behavior (Fig. 1), with a few rabid dogs biting many others (4% of rabid dogs bit >10 other dogs each, and four bit >50 dogs) and running long distances (nine rabid dogs ran >10 km to contact other animals). Overdispersion in the size of transmission chains (fig. S11) reflects this variability in rabid dog behavior and thus makes persistence more remarkable (*13*). Our modeling captured this individual heterogeneity (Fig. 1F), revealing its relevance for rabies dynamics: In counterfactual simulations without individual heterogeneity, rabies incidence was reduced by around 50% (fig. S7C) and the (relatively) large outbreaks observed in nature and in simulations with heterogeneity did not occur (Fig. 3, C versus A). The biological drivers of this variation result predominantly from individual-level (*13*) rather than environmental or population-level differences (*14*, *15*),

such as host density, but are poorly understood. Behavioral manifestations of infection that depend on sociological and pathological factors, such as exposure dose and sites of viral proliferation (*16*), might underpin this individual-level variation.

Both the scale of and heterogeneity in contact and movement are crucial to capturing rabies dynamics (Fig. 3). Density-dependent transmission processes, although well described theoretically (*2*, *17*), are extremely challenging to quantify in relation to spatial scale. This may explain why few empirical studies of directly transmitted diseases have found evidence of strong density-dependent transmission (*18*–*20*). For rabies, the main susceptible depletion mechanisms—deaths of rabid dogs and reexposures of dogs already incubating infection—only curtail transmission if infection is clustered and explain why nonspatial models of rabies do so poorly at replicating dynamics, because rabies incidence is negligible at the population-level. Clustering has been shown to reduce pathogen transmission through build-up of immune individuals (*13*), as well as in the context of redundant biting by insect vectors of disease (*21*); it has also been theorized to reduce transmission in the early stages of epidemics (*22*). For rabies, the incubation period acts in a way similar to that of immunity, resulting in redundant exposures that limit transmission. Natural immunity is not generally considered relevant in canine rabies or required to explain persistence, but antibodies have been detected in healthy unvaccinated dogs (*17*). If short-lived immunity does follow aborted infections, as may be the case for vampire bat rabies (*18*), our expectation is that it would cluster in the same way as incubating infections do, reinforcing local-scale effects. Clustering has been demonstrated to make outbreaks less explosive and to extend persistence (*19*), yet the potential relevance of such microdepletion mechanisms to many pathogens may be underestimated because their measurement relies on sufficiently resolved datasets. Our conclusion, that the relevant spatial scale at which to consider host density is determined by the scale of movements of infectious hosts, offers a starting point for the appropriate spatial scale at which to model other pathogens for which spatially detailed data are lacking.

Our analyses further illustrate the degree to which introduced cases contribute to rabies persistence. In the absence of introductions and under observed levels of vaccination, we expect infection to circulate in the Serengeti district for up to 7 years, typically dying out within 4 years (Fig. 3D). But, with between 8 to 24 rabid dogs arriving each year from neighboring villages, infection persists even under reasonable vaccination coverage, even though most cause only short-lived chains of transmission (fig. S10). In settings where vaccination coverage is negligible (dog populations across much of Africa and Asia), our simulations indicated a mean duration of outbreaks from single introductions of between 10 and 30 weeks; however, the maximum exceeded 12 years (fig. S10). Locally self-limiting clusters of cases recur on the landscape (movie S1) and, in combination with heterogeneous movement and contact, permit the invasion and cocirculation of multiple lineages (Fig. 2) (*20*). Recurrent introductions and extinctions have been reported in many endemic settings (*21*–*23*), and cross-border introductions have led to rabies emergence in several previously rabies-free areas (*24*–*28*). In contrast to diseases such as mosquito-transmitted Dengue (*29*), chains of infection circulate largely independently, given the low prevalence of cases and very localized susceptible depletion. The concurrent extinction of all lineages therefore becomes less probable as more chains of infection cocirculate.

From a practical perspective, our findings explain why dog culling has typically been so ineffective for controlling rabies; dog populations would need to be reduced below very low densities across all areas where infection is circulating. Culling more than 50% of the 400,000 dogs in Flores, Indonesia, had no apparent impact on rabies circulation (*30*). Our results reinforce the message that mass dog vaccination remains the most effective and feasible method of controlling rabies and provides insights that should inform elimination strategies. Simulations indicate that although dog vaccinations prevented more than 4000 animal cases, 2000 human rabies exposures, and 50 deaths in the Serengeti district, introductions continually seeded new foci, with scaled-up dog vaccination (beyond the district) required to achieve elimination (Fig. 3). Infrequent longer-distance movements of rabid dogs seed outbreaks in unaffected localities across heterogeneously distributed populations, leading to localized flare-ups where vaccination coverage is not maintained. The resulting low prevalence persistence presents a challenge for elimination, given that surveillance is very weak in most rabies-endemic regions. Yet the concurrent circulation of viral lineages offers an opportunity for using increasingly affordable genomic approaches to assess the performance of both rabies surveillance and control (*31*). Differentiating undetected circulation from reintroductions will be necessary as control efforts are scaled up toward the 2030 goal of zero human deaths from dog-mediated rabies (*32*).

REFERENCES AND NOTES

1. H. McCallum, N. Barlow, J. Hone, *Trends Ecol. Evol.* **16**, 295–300 (2001).
2. J. Lello, A. Fenton, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **372**, 20160082 (2017).
3. J. Jousimo *et al.*, *Science* **344**, 1289–1293 (2014).
4. R. Anderson, R. May, *Infectious Diseases of Humans: Dynamics and Control* (Oxford Univ. Press, 1991).
5. K. Hampson *et al.*, *PLOS Biol.* **7**, e53 (2009).
6. P. G. Coleman, C. Dye, *Vaccine* **14**, 185–186 (1996).
7. K. Hampson *et al.*, *PLOS Negl. Trop. Dis.* **9**, e0003709 (2015).
8. T. Lembo *et al.*, *J. Appl. Ecol.* **45**, 1246–1257 (2008).
9. Materials and methods are available as supplementary materials.
10. A. M. Czupryna *et al.*, *PLOS ONE* **11**, e0167092 (2016).
11. C. S. Holling, *Can. Entomol.* **91**, 293–320 (1959).
12. J. Antonovics, Y. Iwasa, M. P. Hassell, *Am. Nat.* **145**, 661–675 (1995).
13. J. O. Lloyd-Smith, S. J. Schreiber, P. E. Kopp, W. M. Getz, *Nature* **438**, 355–359 (2005).
14. A. James, J. W. Pitchford, M. J. Plank, *Proc. Biol. Sci.* **274**, 741–747 (2007).
15. O. Tardy, A. Massé, F. Pelletier, D. Fortin, *Ecol. Modell.* **373**, 25–38 (2018).
16. N. Mollentze, D. G. Streicker, P. R. Murcia, K. Hampson, R. Biek, *Proc. Natl. Acad. Sci. U.S.A.* **117**, 28859–28866 (2020).
17. S. Gold, C. A. Donnelly, P. Nouvellet, R. Woodroffe, *PLOS Negl. Trop. Dis.* **14**, e0007933 (2020).
18. J. C. Blackwood, D. G. Streicker, S. Altizer, P. Rohani, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 20837–20842 (2013).
19. T. A. Perkins, T. W. Scott, A. Le Menach, D. L. Smith, *PLOS Comput. Biol.* **9**, e1003327 (2013).
20. K. Brunker *et al.*, *Mol. Ecol.* **27**, 773–788 (2018).
21. H. Bourhy *et al.*, *PLOS Pathog.* **12**, e1005525 (2016).
22. J. Zinsstag *et al.*, *Sci. Transl. Med.* **9**, eaaf6984 (2017).
23. H. L. Beyer *et al.*, *Proc. Biol. Sci.* **278**, 2182–2190 (2011).
24. B. Sharma, N. K. Dhand, N. Timsina, M. P. Ward, *Emerg. Infect. Dis. J.* **16**, 12 (2010).
25. K. Tohma *et al.*, *Infect. Genet. Evol.* **38**, 22–28 (2016).
26. K. Rysava *et al.*, *BMC Infect. Dis.* **20**, 778 (2020).
27. R. Castillo-Neyra *et al.*, *PLOS Negl. Trop. Dis.* **11**, e0005460 (2017).
28. S. E. Townsend *et al.*, *PLOS Negl. Trop. Dis.* **7**, e2372 (2013).
29. H. Salje *et al.*, *Science* **355**, 1302–1306 (2017).
30. C. Windiyaningsih, H. Wilde, F. X. Meslin, T. Suroso, H. S. Widarso, *J. Med. Assoc. Thai.* **87**, 1389–1393 (2004).
31. K. Brunker *et al.*, *Wellcome Open Res.* **5**, 3 (2020).
32. R. Minghui, M. Stone, M. H. Semedo, L. Nel, *Lancet Glob. Health* **6**, e828–e829 (2018).
33. K. Hampson, boydorr/TransmissionScale:. *Zenodo* (2022); doi:10.5281/zenodo.5885134.

ACKNOWLEDGMENTS

We thank local communities and government staff from the animal and public health sectors for ongoing support and the Serengeti Health Initiative for continued dog vaccination. J. Metcalf, D. Streicker, J. Dushoff, M. Li, and three reviewers provided very constructive feedback, which greatly improved the work. We are grateful to MSD Animal Health for donating vaccines for dog vaccination campaigns. **Funding:** This work was supported by Wellcome grants 095787/Z/11/Z and 207569/Z/17/Z (to K.Ha.).
Ethics statement: This study was approved by the Tanzania Commission for Science and Technology, the Institutional Review Boards of the National Institute for Medical Research in Tanzania and of Ifakara Health Institute, and by the Ministry of Regional Administration and Local Government (NIMR/HQ/R.8a/vol.IX/300, NIMR/HQ/R.8a/vol.IX/994, NIMR/HQ/R.8a/vol.IX/2109, NIMR/HQ/R.8a/vol.IX/2788, and IHI/IRB/No:22-2014). **Author contributions:** Conceptualization: K.Ha., R.M., and D.T.H. Methodology: R.M., M.R., E.A.F., and K.Ha. Investigation: K.Ha., M.M., A.L., K.R., M.R., K.B., and K.Ho. Visualization: K.Ha., R.M., M.R., and E.A.F. Funding acquisition: K.Ha. Project administration: K.Ha. and R.K. Supervision: K.Ha. and R.K. Writing, original draft: K.Ha. and R.M. Writing, review and editing: K.Ha., R.M., D.T.H., S.C., K.R., and M.R. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** Code to reproduce the analyses together with deidentified data are available at (*33*).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abn0713
Materials and Methods
Tables S1 and S2
Figs. S1 to S12
References (*34*–*39*)
MDAR Reproducibility Checklist
Movie S1

11 November 2021; accepted 7 March 2022
10.1126/science.abn0713

BATTERIES

Dynamics of particle network in composite battery cathodes

Jizhou Li[1]†, Nikhil Sharma[2]†, Zhisen Jiang[1], Yang Yang[3]‡, Federico Monaco[3], Zhengrui Xu[4], Dong Hou[4], Daniel Ratner[5], Piero Pianetta[1], Peter Cloetens[3], Feng Lin[4]*, Kejie Zhao[2]*, Yijin Liu[1]*

Improving composite battery electrodes requires a delicate control of active materials and electrode formulation. The electrochemically active particles fulfill their role as energy exchange reservoirs through interacting with the surrounding conductive network. We formulate a network evolution model to interpret the regulation and equilibration between electrochemical activity and mechanical damage of these particles. Through statistical analysis of thousands of particles using x-ray phase contrast holotomography in a $LiNi_{0.8}Mn_{0.1}Co_{0.1}O_2$-based cathode, we found that the local network heterogeneity results in asynchronous activities in the early cycles, and subsequently the particle assemblies move toward a synchronous behavior. Our study pinpoints the chemomechanical behavior of individual particles and enables better designs of the conductive network to optimize the utility of all the particles during operation.

Lithium-ion batteries (LIBs), with a high energy density and long lifetime, have been widely adopted for a broad range of applications. The composite cathode of LIBs is made of many electrochemically active particles embedded in a conductive carbon and binder matrix. The microstructure plays a crucial role in governing the LIB performance through modulating the electronic and ionic transport properties (*1*, *2*) and the chemomechanical behavior (*3*–*5*). The cracking, disintegration, and (de)activation behaviors of the electrochemically active cathode particles affect the capacity fade over prolonged battery cycling (*6*–*8*).

Alleviation of the active particle damage has focused on understanding and tuning the morphological and chemical characteristics at the microscale (*9*–*12*), such as particle size, elongation and sphericity (*13*, *14*), crystallographic arrangement (*15*, *16*), mesoscale kinetics (*17*, *18*), grain boundary properties (*19*, *20*), and compositional variation (*21*–*24*). For instance, reducing the primary particle size is an effective approach to improving the fast-charging performance because smaller particles have shorter ion diffusion paths (*25*, *26*). Designing particles with elongated morphology, e.g., in the form of nanoplates or nanorods, can also improve the specific capacity and reduce charge transfer resistance (*27*). However, the correlation between the particle morphology and the cell performance is rather complex, with effects at multiple length and time scales. The dynamics of particle network have substantial impacts but are rarely studied. For example, recent studies have uncovered the local heterogeneity in the electrode, where active particles contribute to the cell-level chemistry differently in time and position (*28*). Some particles release lithium ions at a faster rate than their peers under fast-charging conditions (*29*). Some local regions could become inactive while the cell can still function well as a whole. To make a substantial improvement effectively, the particle structure and the electrode morphology should be tailored coherently, and a synergy could be achieved by doing so. A global homogenization will eventually develop after long-term cycling; however, a poorly designed electrode would reach this state when most of its particles are severely damaged. By contrast, a well-formed electrode would rapidly converge to the electrode synchronization with most of its particles still intact. In our study, we image thick Ni-rich composite cathode electrodes with a multilayer of $LiNi_{0.8}Mn_{0.1}Co_{0.1}O_2$ (NMC) particles at different states using nano-resolution hard x-ray phase contrast holotomography (Fig. 1A). These electrodes are recovered from standard coin cells that were cycled under fast-charging conditions for 10 cycles and 50 cycles (fig. S1), respectively. With high spatial resolution and contrast, and a large field of view, our three-dimensional imaging data cover a large number of active particles that demonstrate a wide variety of damage patterns. To facilitate a statistical analysis, we build on our previous neural network–based particle identification method (*1*) and improve its accuracy and efficiency by developing a diagonal data fusion approach, which is illustrated in figs. S2 and S3. Upon completion of the particle identification using this method, the damage level of individual particles is further quantified. The relative probability distribution of the particle damage degree is presented in Fig. 1B, and a few randomly

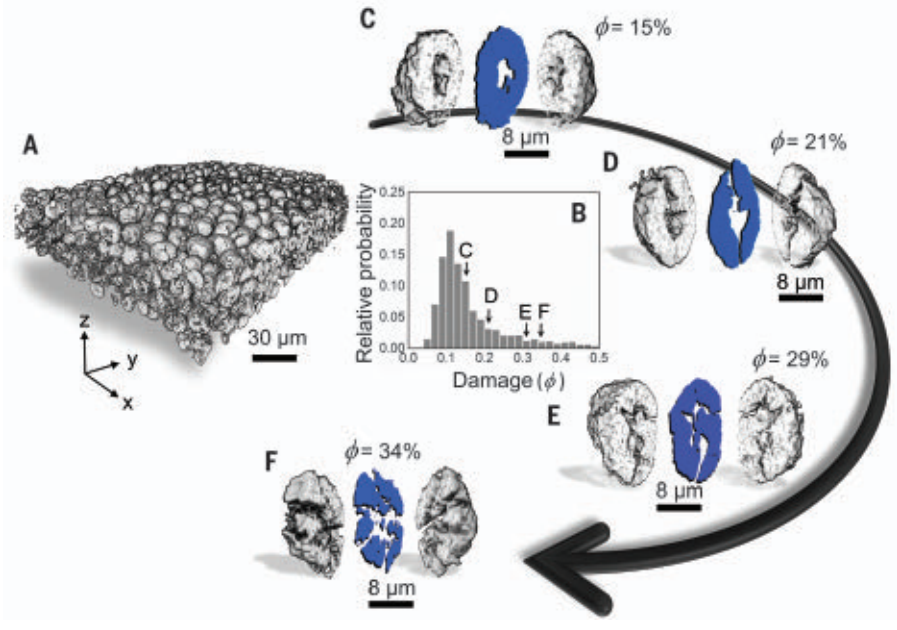


Fig. 1. Imaging cathode electrodes with a multilayer of NMC particles using nano-holotomography. (**A**) Visualization of the composite battery cathode obtained by synchrotron nano-holotomography. Each NMC particle has its own properties in position, particle structure, mesoscale chemistry, and local morphology. (**B**) Probability distribution of the particle damage level. (**C** to **F**) Randomly selected examples of NMC particles with different levels of damage.

[1]Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA. [2]School of Mechanical Engineering, Purdue University, West Lafayette, IN 47906, USA. [3]European Synchrotron Radiation Facility, Grenoble 38000, France. [4]Department of Chemistry, Virginia Tech, Blacksburg, VA 24061, USA. [5]Machine Learning Initiative, SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA.
***Corresponding author. Email: fenglin@vt.edu (F.L.); kjzhao@purdue.edu (K.Z.); liuyijin@slac.stanford.edu (Y.L.)**
†These authors contributed equally to this work.
‡Present address: National Synchrotron Light Source II, Brookhaven National Laboratory, Upton, NY 11973, USA.

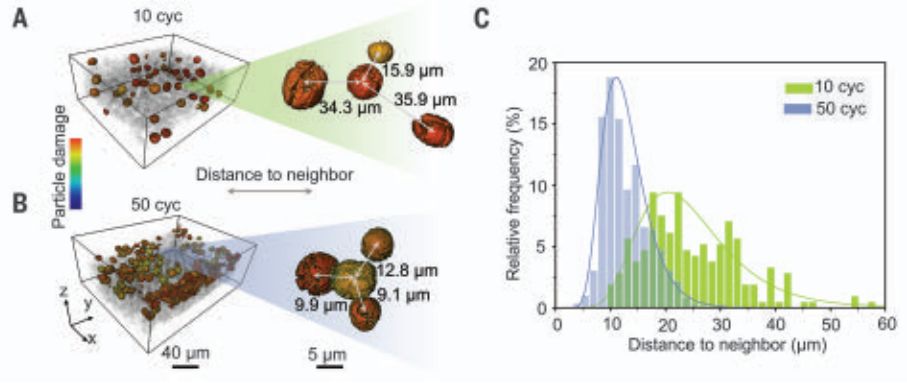


Fig. 2. Heterogeneous particle damage in battery electrodes. The spatial distributions of the severely damaged particles in the (**A**) 10-cycled and (**B**) 50-cycled electrodes. The degree of particle damage is color coded. Selected representative regions are enlarged for better visualization. The distances between the central damaged particle and its three nearest-neighboring damaged particles are annotated in the enlarged view. (**C**) Probability distributions of the distance between two neighboring severely damaged particles in 10-cycled and 50-cycled electrodes.

selected particles with different damage patterns are highlighted in Fig. 1, C to F.

The severely damaged particles are those overused during the electrochemical fast-charging process. Their spatial distribution and arrangement are evidence of the spatially heterogeneous electrode utility. As shown in Fig. 2, A and B, the severely damaged particles are sparsely distributed in the 10-cycled electrode and, their concentration increased upon further cycling, featuring a denser agglomeration in the 50-cycled electrode (as illustrated in the enlarged views). Figure 2C shows the probability distributions of the distance between two neighboring severely damaged particles in 10-cycled and 50-cycled electrodes, respectively. A shift toward shorter distance can be observed in the 50-cycled electrode, indicating a synchronization effect within the local particle clusters.

We perform theoretical modeling to understand the damage and Li reaction behaviors of NMC particles across cycles. We envision that the interplay between the electrochemical activity and the mechanical damage regulates the performance of the NMC particles. The deeper state of charge incurs more severe consequences, such as the particle damage and decohesion of particles from the conductive agent. By contrast, the mechanical damage increases the cost of Li reactions and suppresses the electrochemical activity of individual particles. A more synchronous behavior of the composite cathode is achieved in the prolonged cycles because of the equilibrium between the electrochemical activity and mechanical damage. To test this hypothesis, we conducted finite element analysis to model the electrochemical response and mechanical damage of a NMC cathode composed of three spherical NMC active particles surrounded by two homogeneous porous carbon/binder (CB) domains of different electrical conductivities (Fig. 3). The intention of the computational model is not to capture all the explicit microstructural details in the composite cathode. Rather, our goal is to replicate the salient feature in the composite such that the active NMC particles are covered by different degrees of the electrically conductive agent, which results in various local conducting networks for individual particles. In this simplified model, the surrounding high-conductivity and low-conductivity CBs differ in their electrical conductivities. The model assumes that liquid electrolyte is soaked in the porous CB domains and thus the NMC particles are fully accessible to the Li^+ in the liquid electrolyte. We set different ratios of coverage of CB on the periphery of each active particle, as illustrated in Fig. 3A. The interface of active particles attached to the high-conductivity CB undergoes a faster electrochemical reaction than the boundary enclosed by the low-conductivity CB. Thus, each active particle experiences dissimilar electrochemical activities, as inferred from the diverging concentration profiles (C/C_{max}) during the first charging process (Fig. 3B).

The modulation between the electrochemical activity and mechanical damage reduces the variation of Li concentration with the progression of (dis)charge cycles. As shown in the normalized Li concentration plot (Fig. 3B) and the plot of the Li concentration variation across the three NMC particles (Fig. 3C), the concentration profiles converge with battery operation. During the charging process, Li extraction generates a reduction of the lattice volume in the NMC particles (*30*). Considerable variation in the mechanical properties of the NMC particles (elastic modulus ~ 140 GPa) and CB (elastic modulus ~ 2 GPa) generates strain mismatch at the interface. The apex of the mismatch occurs near the end of the charging process, as demonstrated by the divergence of the damage profiles in Fig. 3D and their corresponding differences to the mean damage in Fig. 3E.

After the onset of heterogeneous damage among the NMC particles, the individual damage curves diverge (Fig. 3D). The individual NMC particle characteristic—i.e., the dissimilar coverage by high- and low-conductivity CBs—commands the degree of heterogeneous damage in the early cycles. With successive discharging and charging processes, the modulation between electrochemical activity and mechanical damage reduces the imbalance within the system (through the interfacial resistance for charge transfer). Consequently, the damage level for all three particles converges, demonstrating the system's progression toward a synchronized behavior. In addition, we observe a similar transformation to synchronized damage behavior for the system with more NMC active particles (fig. S4). The variation in the periphery contact with high- and low-conductivity CB regions generates heterogeneous reactions for each particle. After the initial divergence, the individual particle damage is tuned by the feedback to electrochemical activity that progresses toward a synchronized behavior in the composite electrode.

Both the particle damage and Li concentration profiles theoretically confirm the asynchronous to synchronous evolution in composite cathodes. Such a transition can occur for a number of reasons, such as from the particles' self-attributes, interactions with neighboring particles, and CB domains. To probe the evolution mechanisms from the intrinsic or internal physical nature of cathode particles, we analyzed the three-dimensional tomographic imaging data through an interpretable machine learning framework.

Using more than 2000 accurately identified NMC particles, we extract their structural, chemical, and morphological characteristics. More specifically, we divided the particle attributes into four different groups: position, chemical properties, particle structure, and local morphology (as illustrated in table S1 and fig. S5; in total, 24 attributes are extracted). As indicated in fig. S6, these extracted particle attributes demonstrate varying characteristics in their respective probability distributions. Revealing their interrelationship is not straightforward and could benefit from more advanced computing and modeling approaches.

We leverage the advances in machine learning to model relationship and dependencies among attributes, i.e., descriptors of the cathode particles' properties (fig. S7). The model has to be both accurate and interpretable. To elucidate the intertwined limiting factors for battery cathode robustness, we explore the degree of

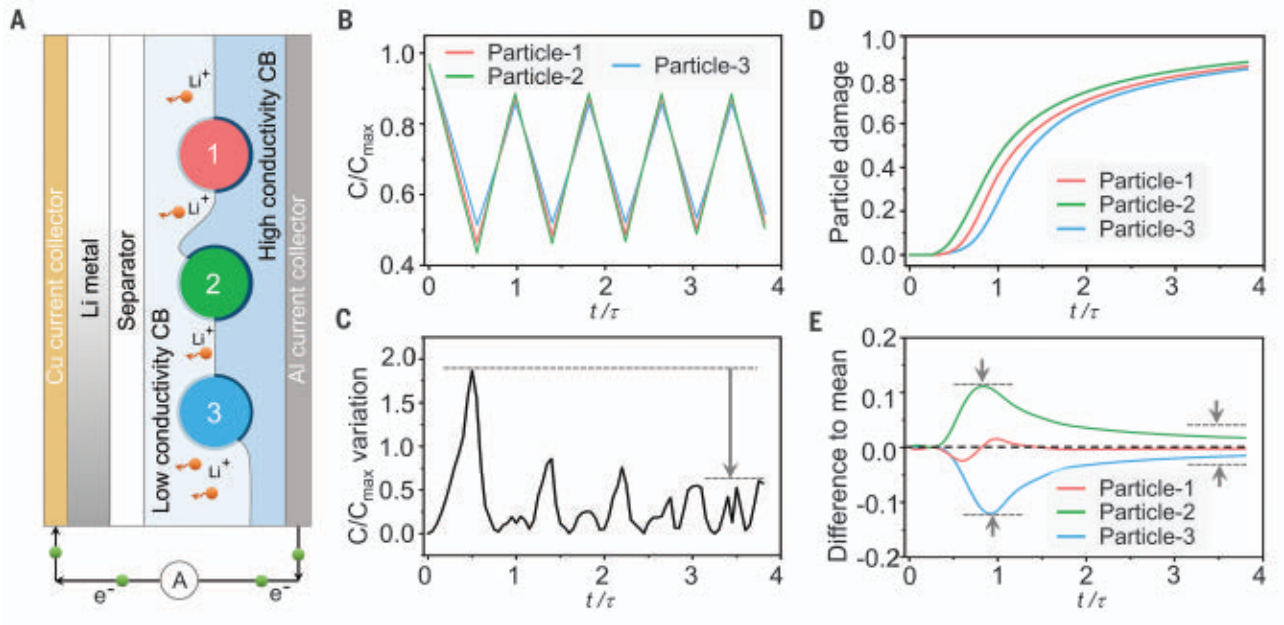


Fig. 3. Finite element analysis of the electrochemical activity and mechanical damage in the NMC cathode. (**A**) Illustration of the composite model during the charging process in the battery. (**B**) Normalized Li concentration profiles depict the inherent heterogeneity of the system during the first charging process with respect to the normalized time t/τ, where t is the real time in Li reactions and τ = 720 s is the theoretical time to reach the full capacity of NMC. Although the particles start with the same state of charge, Li concentration differs at the end of the first charge process. (**C**) The variation of Li concentration profiles among three NMC particles. The overall trend demonstrates the tendency toward a synchronized behavior. (**D**) The damage profiles for three NMC active particles diverge near the end of the first charge process. With the progression of the cycling process, the damage profiles for all three particles converge. (**E**) Each particle's deviation from the mean damage profile (the black dashed line).

cathode NMC particles' engagement in the cell-level chemistry through attribute correlation and damage regression. These two steps are accomplished using a regularized autoencoder neural network (*31*) and random forest (RF) regression (*32*). The SHAP (Shapley additive explanations) (*33*) is utilized to rank the importance of the particle properties to the degree of particle damage during the process of regression, which effectively reveals the contributions of different microstructural characteristics to the damage profile for every particle in our electrode. The circular plot (*34*) is used for better visualization of the Pearson's correction among different particle attributes. Integrated with the SHAP values, the RF provides not only accurate regression results, but also the interpretability of the impacts of all the input attributes on individual predictions as well as global insights.

Specifically, the autoencoder neural network compresses the input attributes into latent dimension (LDs) through an encoder network (fig. S8), which has been extensively applied for scientific discovery because of its ability to learn nonlinear functions and its good interpretability (*35*, *36*). The LDs of both datasets for 10-cycled and 50-cycled electrodes, respectively, are calculated and are subsequently correlated with each other through Pearson's correlation. As shown in fig. S9, different LDs show intertwined relationships in both cases. Each node in the circular plot represents one LD, and a connection between two nodes indicates a relatively high correlation between them. The sign of the correlation coefficient (+/−) defines the direction of the relationship. For the 10-cycled electrode, the first five LDs appear to be independent. As more LDs are added, we start to observe correlations among them. For the 50-cycled electrode, in addition to the common connections, several additional correlations emerge (as highlighted by the dark colors in fig. S9D). The observation of a higher degree of interdependence among different LDs indicates that the particles' structural and chemical characteristics become more intertwined upon battery cycling.

When interpreting the model regression results with the SHAP values, which utilize the game-theory–based Shapley values (*37*), the contribution of each attribute to the model's output (the particle damage degree) can be obtained. Attributes with larger SHAP values are considered to be more important to the target damage degree. The contribution scores of all attributes to the particle damage in both 10-cycled and 50-cycled electrodes are presented in Fig. 4 (attributes are grouped on the basis of their properties and reordered for better visualization), and the interpretation is provided below. The advantage of using SHAP to explain the regression model is its superior robustness to correlated attributes compared to the traditional methods (*38*), e.g., the Pearson's correlation, which cannot systematically capture the key differences in the studied electrodes (fig. S11).

From our model-based prediction in Fig. 4, some of the attributes follow expected trends. For example, the particle's depth, *Z*, affects the particle damage (*28*, *39*). This can be related to the cell polarization effect, which results in particles at different depths experiencing different states of charge at a given time. The *Z*-dependence of particle damage is more pronounced in the 10-cycled electrode, in good agreement with previous reports (*9*, *28*). The electron density, *EDensity*, has been associated with the state of charge (*1*) and its averaged value and degree of variation, *Homogeneity*, shows considerable impact on the particle damage throughout the cycling process. The surface area and roughness (*RoughOuter*, *RoughInner*, *SurfOuter*, and *SurfInner*; table S1) could affect the cohesion of the active particles and the CB matrix. Therefore, the surface characteristics could affect the particle damage. The particle's size, *Volume*, appears to be correlated with the particle damage. Its contribution score seems slightly lower in the 50-cycled electrode. This trend suggests that the particle-size effect might

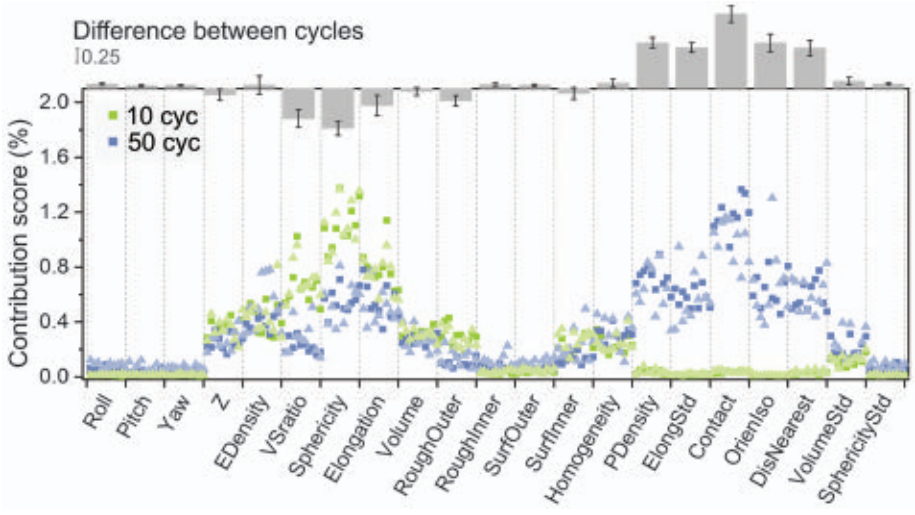


Fig. 4. Interpretable machine learning framework for particle attributes modeling. The contribution scores of all attributes to the particle damage in 10-cycled (green) and 50-cycled (blue) electrodes. Triangle and square markers represent results from two robustness validation approaches, data-subsampling and random-seeding, respectively. The mean and SD of the differences in the contribution scores (N = 20) between the 10-cycled and 50-cycled data are plotted on the top.

not be the limiting factor in the later cycles. However, the variation of the neighboring particles' volumes, *VolumeStd*, shows an opposite trend, featuring a higher contribution score in the 50-cycled electrode. This observation indicates that, upon prolonged battery operation, the uniformity in the neighboring particle size becomes a more relevant factor that affects the particle damage. Mixing particles of different sizes has been used as a method to improve the electrode's packing density. Our result suggests that this approach should be carefully examined from the long-term cyclability perspective. Another finding is that the particle elongation, *Elongation*, which represents the ratio of the longest axis length to the shortest axis length (table S1), has a rather high contribution score, which, however, decreases upon cycling. By contrast, the alignment of the neighboring particles, *OrienIso*, which shows negligible contribution in the 10-cycled electrode, becomes more important in the later cycles. In advanced battery electrode manufacturing, particle alignment can be purposely adjusted by controlling externally applied electric and/or magnetic fields (*40*–*42*).

When visualizing the overall picture of our statistical analysis over thousands of particles, we find an interesting pattern: In the early cycles, individual particles' characteristics (e.g., the position *Z*, *VSratio*, *Sphericity*, and *Elongation*; table S1) predominantly determine their respective degrees of damage, featuring an asynchronous behavior that is in agreement with our theoretical modeling result. In the later cycles, however, the interplay among neighboring particles (e.g., *Contact*, *DisNearest*, *OrienIso*, and *PDensity*) becomes more important, which indicates that the local interparticle arrangement can critically affect the asynchronous-to-synchronous transition. The mean difference and standard deviation (SD) of each attribute's contribution scores between the 10-cycled and 50-cycled datasets is presented on the top of Fig. 4, featuring a valley on the left and a peak on the right, supporting the above-described observation.

Our experimental observations (Fig. 2) and machine learning analysis (Fig. 4) collectively corroborate the theoretical modeling (Fig. 3). These results reveal a transition from the asynchronous behavior in the early stage toward a synchronous state later in the particle network evolution, where the interplay among neighbor particles plays a facilitating role (fig. S12). Particles' self-attributes, together with the dynamic nature of the conductive network, jointly determine the damage behavior of NMC particles in composite electrodes. These are critical factors for cathode design to prolong the cycle life of batteries. On the basis of our results, in the active cathode powder, it is useful to suppress the particle-to-particle variation in their structural characteristics, such as particle size, sphericity, elongation, etc. At the electrode scale, an ordered particle arrangement is favorable, which can be reinforced through a field-guided approach. Whereas the in-plane homogeneity is desirable, in the out-of-plane direction, a structural gradient could be beneficial because of the electrochemical polarization, which is more severe in thick electrodes. To summarize, an ordered electrode configuration with tailored depth-dependent packing of uniform active particles would be robust to prolonged battery cycling.

From the synthesis perspective, the particle shape and structure can be tuned by controlling the sintering temperature, incorporation of trace-element doping, designing the architecture of the precursor, and surface coating. These are common synthesis strategies and can be scalable for mass production. For the electrode manufacturing, the field-guided approach has been demonstrated to be effective for creating an ordered structure. This is compatible with the existing electrode manufacturing facilities and, thus, can be fairly cost effective.

REFERENCES AND NOTES

1. Z. Jiang *et al*., *Nat. Commun.* **11**, 2310 (2020).
2. D. E. Stephenson *et al*., *J. Electrochem. Soc.* **158**, A781 (2011).
3. R. Xu *et al*., *J. Mech. Phys. Solids* **129**, 160–183 (2019).
4. Z. Xu *et al*., *J. Mater. Chem. A* **6**, 21859–21884 (2018).
5. L. Mu *et al*., *Nano Lett.* **18**, 3241–3249 (2018).
6. Z. Yang, F. Lin, *J. Phys. Chem. C* **125**, 9618–9629 (2021).
7. H. Liu *et al*., *Nano Lett.* **17**, 3452–3457 (2017).
8. Y. Mao *et al*., *Adv. Funct. Mater.* **29**, 1900247 (2019).
9. F. Lin, K. Zhao, Y. Liu, *ACS Energy Lett.* **6**, 4065–4070 (2021).
10. H. H. Sun *et al*., *ACS Energy Lett.* **5**, 1136–1146 (2020).
11. W. E. Gent *et al*., *Adv. Mater.* **28**, 6631–6638 (2016).
12. X. Lu *et al*., *Nat. Commun.* **11**, 2079 (2020).
13. X. Liu *et al*., *Adv. Energy Mater.* **11**, 2003583 (2021).
14. W. Li, E. M. Erickson, A. Manthiram, *Nat. Energy* **5**, 26–34 (2020).
15. Z. Xu *et al*., *Nat. Commun.* **11**, 83 (2020).
16. A. Liu *et al*., *J. Electrochem. Soc.* **168**, 050506 (2021).
17. A. J. Merryweather, C. Schnedermann, Q. Jacquet, C. P. Grey, A. Rao, *Nature* **594**, 522–528 (2021).
18. S. Kuppan, Y. Xu, Y. Liu, G. Chen, *Nat. Commun.* **8**, 14309 (2017).
19. M. Yoon *et al*., *Nat. Energy* **6**, 362–371 (2021).
20. J.-N. Zhang *et al*., *Nat. Energy* **4**, 594–603 (2019).
21. Y.-K. Sun *et al*., *Nat. Mater.* **8**, 320–324 (2009).
22. F. Lin *et al*., *Nat. Energy* **1**, 15004 (2016).
23. Z. Zhu *et al*., *Nat. Energy* **4**, 1049–1058 (2019).
24. H. H. Sun *et al*., *Nat. Commun.* **12**, 6552 (2021).
25. C. Delacourt, P. Poizot, S. Levasseur, C. Masquelier, *Electrochem. Solid-State Lett.* **9**, A352 (2006).
26. L. Bläubaum *et al*., *ChemElectroChem* **7**, 4755–4766 (2020).
27. D.-H. Kim, J. Kim, *J. Phys. Chem. Solids* **68**, 734–737 (2007).
28. Y. Yang *et al*., *Adv. Energy Mater.* **9**, 1900674 (2019).
29. J. Park *et al*., *Nat. Mater.* **20**, 991–999 (2021).
30. J. Li *et al*., *Sustain Mater Technologies.* **29**, e00305 (2021).
31. M. A. Kramer, *AIChE J.* **37**, 233–243 (1991).
32. L. Breiman, *Mach. Learn.* **45**, 5–32 (2001).
33. S. M. Lundberg *et al*., *Nat. Mach. Intell.* **2**, 56–67 (2020).
34. M. Krzywinski *et al*., *Genome Res.* **19**, 1639–1645 (2009).
35. R. Gómez-Bombarelli *et al*., *ACS Cent. Sci.* **4**, 268–276 (2018).
36. G. Qian *et al*., *ACS Energy Lett.* **6**, 687–693 (2021).
37. L. S. Shapley, "A value for n-person games" (RAND Corporation, 1953); https://www.rand.org/pubs/papers/P295.html.
38. S. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions arXiv:1705.07874 [cs.AI] (2017).
39. J. Hu *et al*., *J. Power Sources* **454**, 227966 (2020).
40. J. Billaud, F. Bouville, T. Magrini, C. Villevieille, A. R. Studart, *Nat. Energy* **1**, 16097 (2016).
41. J. S. Sander, R. M. Erb, L. Li, A. Gurijala, Y.-M. Chiang, *Nat. Energy* **1**, 16099 (2016).
42. C. Lei, Z. Xie, K. Wu, Q. Fu, *Adv. Mater.* **33**, e2103495 (2021).
43. J. Li *et al*., Data for: Dynamics of particle network in composite battery cathodes, Zenodo (2022); https://doi.org/10.5281/zenodo.5888945.

ACKNOWLEDGMENTS

Funding: The work at SLAC National Accelerator Laboratory is supported by the Department of Energy, Laboratory

Directed Research and Development program under contract DE-AC02–76SF00515. The work at Purdue University is supported by the National Science Foundation under grant no. DMR-1832707. The work at Virginia Tech is supported by the National Science Foundation under grant no. DMR-1832613. **Author contributions:** Conceptualization: Y.L., K.Z., and F.L.; Investigation: J.L., N.S., Z.J., Y. Y., F. M., Z. X., D.H., D.R., P.P., P.C., F.L., K.Z., and Y.L.; Methodology: J.L., N.S., F.L., K.Z., and Y.L.; Resources: F.L., K.Z., and Y.L.; Writing – original draft: J.L., N.S., K.Z., F.L., and Y.L. All authors reviewed, edited, and approved the manuscript. J.L. and N.S. contributed equally to this work. **Competing interests:** None declared. **Data and materials availability:** The source code supporting the findings of the study is available at (*43*). All (other) data needed to evaluate the conclusions in the paper are present in the paper or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abm8962
Materials and Methods
Figs. S1 to S12
Table S1
References (*44*–*48*)

19 October 2021; accepted 25 February 2022
10.1126/science.abm8962

GRAVITATIONAL WAVES

A gamma-ray pulsar timing array constrains the nanohertz gravitational wave background

The Fermi-LAT Collaboration*†

After large galaxies merge, their central supermassive black holes are expected to form binary systems. Their orbital motion should generate a gravitational wave background (GWB) at nanohertz frequencies. Searches for this background use pulsar timing arrays, which perform long-term monitoring of millisecond pulsars at radio wavelengths. We used 12.5 years of Fermi Large Area Telescope data to form a gamma-ray pulsar timing array. Results from 35 bright gamma-ray pulsars place a 95% credible limit on the GWB characteristic strain of 1.0×10^{-14} at a frequency of 1 year^{-1}. The sensitivity is expected to scale with t_{obs}, the observing time span, as $t_{obs}^{-13/6}$. This direct measurement provides an independent probe of the GWB while offering a check on radio noise models.

Pulsars are spinning neutron stars that emit beams of broadband radiation from radio to gamma-ray wavelengths that appear to pulse as they periodically sweep across the line of sight to Earth (*1*). Millisecond pulsars (MSPs) spin at hundreds of hertz and pulse with sufficient regularity to function as celestial clocks, distributed across the sky and throughout the Galaxy. Timing of individual MSPs by use of radio telescopes has been used to test general relativity and alternative theories of gravity (*2*). Long-term monitoring campaigns of ensembles of MSPs are used to search for low-frequency gravitational waves (GWs), which are expected to be emitted by supermassive black hole (SMBH) binaries that are predicted to exist at the centers of galaxies that have undergone mergers. General relativity predicts that a circular binary with orbital frequency $f/2$ will emit GWs with frequency f and amplitude $\propto f^{2/3}$ (*3*). When SMBH binaries have an orbital separation of ~0.01 pc, which is equivalent to ~2000 astronomical units, the orbits decay primarily through GW emission. Because of this link between GW frequency and amplitude, the superposition of GWs from many SMBH binaries throughout the Universe is predicted to build up a GW background (GWB) with a characteristic GW strain h_c following a power law in frequency (*4*)

$$h_c(f) = A_{gwb}\left(\frac{f}{\text{year}^{-1}}\right)^{\alpha} \quad (1)$$

The spectral index α is predicted to be $-2/3$ for GW-driven binary inspirals, and the dimensionless strain amplitude A_{gwb} incorporates the growth, masses, and merger rates of SMBHs. If SMBHs do not rapidly migrate to the centers of newly merged galaxies, there will be fewer wide binaries, reducing the GW power at low frequencies. Thus, the measured GWB is expected to carry information about the distribution of SMBH masses and the dynamical evolution of SMBH binary systems (*5*).

Searches for the GWB can be performed with ensembles of MSPs—known as pulsar timing arrays (PTAs) (*6*, *7*)—by monitoring the times of arrival (TOAs) of the steady pulses from each pulsar, which arrive earlier or later than expected owing to the spacetime perturbations. Because the GWB is expected to be a sum of many individual sources, the induced TOA variations are random and differ for each pulsar but have a common spectrum of power spectral densities, $P(f)$

$$P(f) = \frac{A_{gwb}^2}{12\pi^2}\left(\frac{f}{\text{year}^{-1}}\right)^{-\Gamma}\text{year}^{-3} \quad (2)$$

with spectral index $\Gamma = 3 - 2\alpha = 13/3$ for SMBHs (*4*). This functional form has more power at low frequencies so is referred to as a red spectrum. For observations taken at an approximately fixed location (Earth), the GWB is expected to produce a signature quadrupolar pattern of TOA variations, known as the Hellings-Downs correlation (*8*).

Because the expected quadrupolar correlations are only about 10% of the total signal, the GWB is predicted to initially appear as a set of independent signals from each pulsar, with power spectra all consistent with Eq. 2. The quadrupolar distribution would only become evident in more sensitive observations. Radio PTAs have reported a red spectrum process with modest statistical significance (*9*–*12*), but no Hellings-Downs correlation has been found. These results could be compatible with $\alpha = -2/3$ and $A_{gwb} \sim 2 \times 10^{-15}$ to 3×10^{-15} at 1 year^{-1} (Fig. 1). This would be consistent with some predictions for the GWB (*5*), but because no spatial correlations have been detected, it could have other origins.

A potential alternative explanation for this signal is spin noise, which is approximately power-law red noise intrinsic to each pulsar, with some MSPs observed to have a spin noise spectral index (Γ) of 2 to 7 (*13*, *14*). Possible physical origins for spin noise include turbulence in the neutron star interior (*15*) and systematic variations in the magnetic field and corotating plasma, which govern the rotational energy loss of the pulsar (*16*). Pulsars that have spin noise spectra with similar shapes but different amplitudes—which is inconsistent with a GWB—could masquerade as a common mode signal without a Hellings-Downs correlation (*10*).

Another potential noise source for radio PTAs is the frequency-dependent effect of radio propagation through plasma, including the solar wind and the ionized interstellar medium (IISM). Pulsed radio emission at frequency ν is delayed by time τ_{DM}

$$\tau_{DM} = 4.15\ \text{ms} \times \left(\frac{\text{DM}\ \text{cm}^3}{\text{pc}}\right) \times \left(\frac{\nu}{\text{GHz}}\right)^{-2} \quad (3)$$

where DM is the dispersion measure, equal to the total electron column density. The DM of a pulsar can vary with time because of the relative motions of Earth and the pulsar. Correcting for this effect requires repeated measurements by use of multifrequency radio observations and the introduction of many additional degrees of freedom to timing models. Because the propagation paths of radio waves

*Fermi-LAT Collaboration authors and affiliations are listed in the supplementary materials.
†Corresponding authors: Matthew Kerr (matthew.kerr@gmail.com); Aditya Parthasarathy (adityapartha3112@gmail.com)

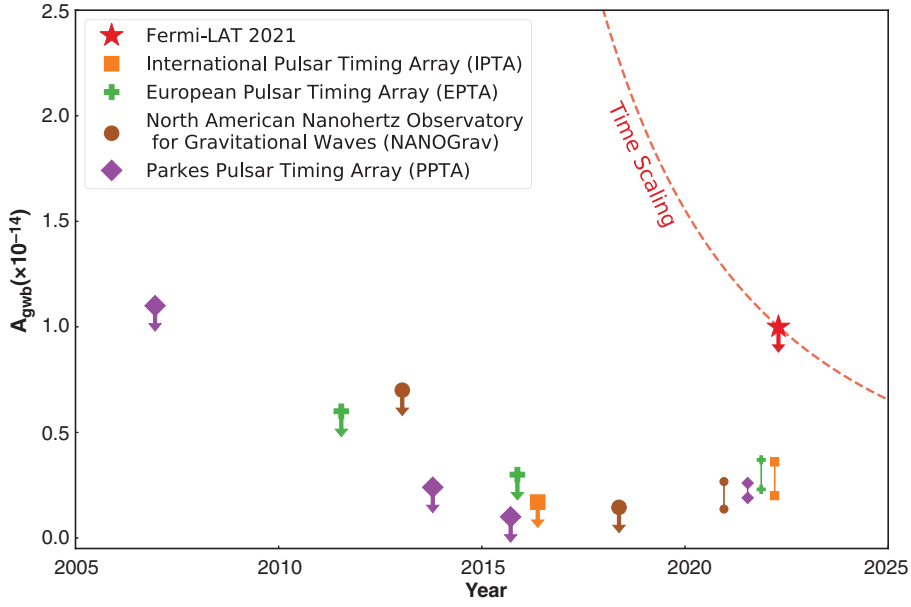


Fig. 1. Constraints on the GW background from radio and gamma-ray PTAs. The inferred constraints on the GWB amplitude at 1 year^{-1} (A_{gwb}) are plotted as a function of publication date (data sources are listed in table S7) and assume $\alpha = -2/3$, as predicted for the superposition of GWs from merging black holes. Colored symbols correspond to each of the PTAs indicated in the key. Upper limits at 95% confidence are shown as downward arrows, and amplitude ranges indicate detections of a common noise process, which could be the GWB or have other origins. The Fermi-LAT 95% upper limit, 1.0×10^{-14}, uses data obtained up to January 2021 and is plotted at a publication date of April 2022. The dashed red line indicates the expected scaling of the Fermi-LAT limit as a function of time.

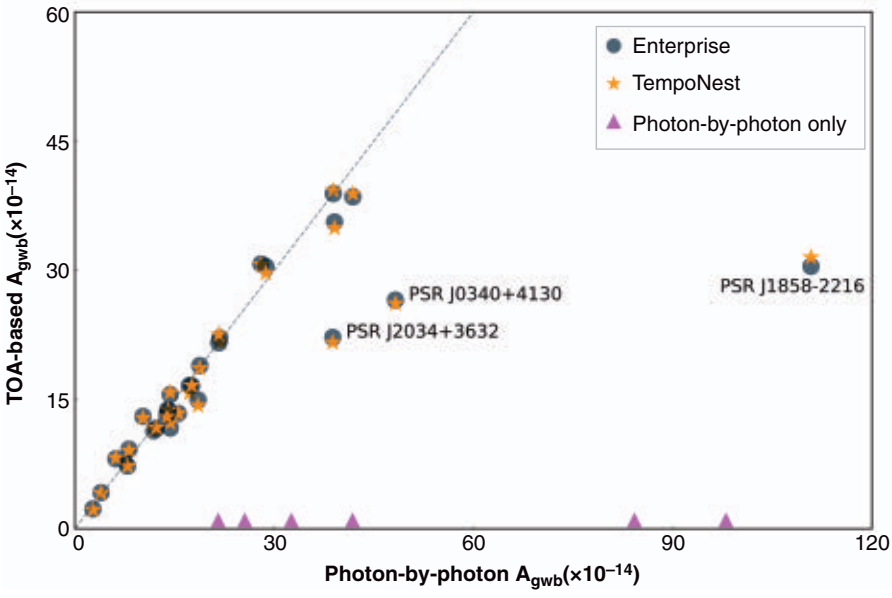


Fig. 2. Comparison between A_{gwb} measurements from each pulsar by using three analysis methods. Data points indicate the limits on an $\alpha = -2/3$ GWB for 35 MSPs computed with three methods: Two TOA-based codes, TEMPONEST (orange stars) and ENTERPRISE (gray circles), are shown as a function of the limit from a photon-by-photon analysis (*x* axis). The dashed line indicates equality between the results of the TOA-based and photon-by-photon methods. Six pulsars (purple triangles) have only a photon-based analysis so are plotted arbitrarily at zero on the *y* axis. The three labeled pulsars are outliers (*21*).

through the IISM depend on ν, the DM itself is frequency dependent (*17*, *18*), so some of this delay is intrinsically unmeasurable. Other propagation effects include a broadening of the pulse, which can only be corrected for bright pulsars, with some components also being unmeasurable (*19*). Because the IISM is turbulent, these uncorrected delays introduce additional red noise to radio pulsar timing data. The variable solar wind introduces similar dispersive delays that can in principle be measured like DM variations but are only partially included in current models (*20*). Because of the wide angular extent of the solar wind, uncorrected delays would be correlated among pulsars. As with spin noise, IISM-induced noise with similar spectra could mimic a GWB signal. Predicted noise amplitudes are similar to the expected GWB signal, but these predictions rely on assumptions about the turbulent spectra of the IISM, which are poorly constrained by data (*19*). Further discussion of the modeling and impact of noise is available in (*21*).

Gamma-ray observations offer a potentially complementary approach: The much higher photon frequency means that the effects of the IISM and solar wind are negligible. The Large Area Telescope (LAT) (*22*), on the Fermi Gamma-ray Space Telescope, is sensitive to giga–electron volt gamma-ray photons emitted by MSPs. Its 2.4-sr field of view performs a continuous survey, covering the full sky every two orbits (~3 hours). Its GPS clock records photon arrival times with <300 ns precision (*23*), enabling pulsar timing. Analyses of LAT survey data have detected 127 of the more than 400 known MSPs in the Milky Way (*21*, *24*). The number of MSPs in this sample, long observing span, and instrumental stability enable a gamma-ray PTA whose characterization of spin noise and a potential GWB signal is free from IISM effects.

Using the 35 brightest and most stable gamma-ray MSPs and 12.5 years of Fermi-LAT data, we searched for the GWB using two different techniques (*21*). First, we implemented a coherent photon-by-photon analysis that retains <1 μs resolution. Second, for analysis with established software used for radio PTAs, we directly measured TOAs from the LAT data (*25*). Because the TOA estimation procedure requires averaging up to 1 year of data, this method loses sensitivity to signals with shorter time scales, and only 29 of the 35 pulsars are suitable.

For each pulsar, we searched for spin noise and derived an upper limit on A_{gwb} using (i) the photon-by-photon method and (ii) two TOA-based software packages, TEMPONEST (*26*) and ENTERPRISE (*27*). None of the pulsars show evidence for spin noise (*21*), and the three different methods provide consistent results for each pulsar (Fig. 2), except in three cases (*21*).

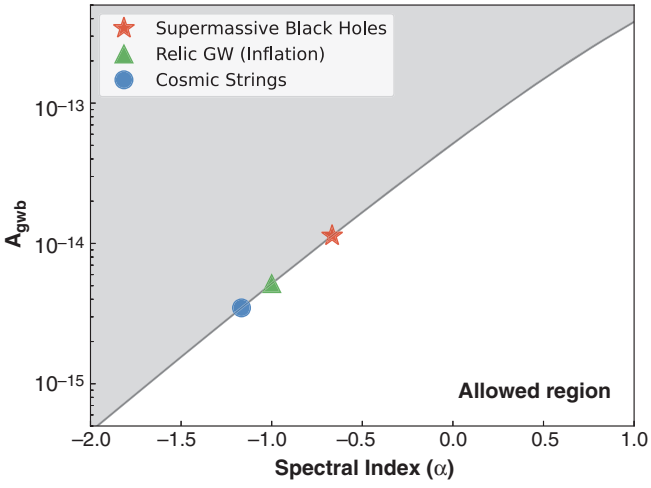


Fig. 3. Gamma-ray constraints on different types of GWB sources. GWB amplitudes A_{gwb} for assumed spectral indices α in the shaded region are excluded with 95% confidence. The symbols indicate the values of α expected for SMBH binaries (red star; our fiducial result), GWs generated during cosmic inflation (green triangle), and from hypothetical cosmic strings (blue circle).

Three of the pulsars in our sample have spin noise measurements from radio PTAs. Using the power spectral indices Γ measured from the radio timing data, we calculated 95% upper limits on spin noise amplitudes from the gamma-ray data. Our limits are below the previously measured values for PSR J0030+0451 (10% of the measured value) and PSR J1939+2134 (60 to 70%) but are unconstraining for PSR J0613−0200. This discrepancy might indicate contamination by residual IISM effects on the radio-based spin noise and GWB signal measurements. We combined the single pulsars into a PTA and estimated A_{gwb} limits under a variety of scenarios, including marginalization over possible spin noise and uncertainties in the position of Earth relative to the Solar System barycenter, and both excluding and including the expected Hellings-Downs quadrupolar spatial correlations (*21*). The resulting representative 95% confidence limit is $A_{gwb} < 1.0 \times 10^{-14}$ (Fig. 1), a factor of 3 to 5 greater than the red spectrum process detected by radio PTAs.

For an idealized PTA, when a potential GWB signal is weak compared with other noise, the signal-to-noise ratio grows proportionally to $A_{gwb}^2 \times t_{obs}^{\Gamma}$ (*28*, *29*), where t_{obs} is the observing time span and Γ = 13/3 for SMBHs, as in Eq. 2. This means that upper limits on A_{gwb} improve following the relation $A_{gwb} \propto t_{obs}^{-13/6}$. However, if the signal detected by radio PTAs does arise from the GWB, then these PTAs are now in the strong signal regime, and their sensitivity will improve more slowly ($\propto t_{obs}^{-1/2}$). The differing time scalings and noise sources allow the gamma-ray PTA data to distinguish residual IISM variations from a potential GWB signal.

The Fermi PTA data have an essentially constant experimental setup; the data are almost uninterrupted, and calibrations have been constant for the full 12.5-year dataset. Gamma-ray data are potentially less subject to astrophysical effects, such as changes in the radio pulse shape (*21*). This stability is particularly useful for probing GWs with frequencies below 0.1 $year^{-1}$. Such low frequencies are predicted to constrain the spectral shape of the GWB, which contains information about the physical sources (*5*).

There are other potential sources of power-law GWBs with different spectral indices, α, such as α = −1 for relic GWs originating during scale-invariant inflation in the early Universe (*30*). Decay of (hypothetical) cosmic strings could also produce power-law spectra under a variety of scenarios (*31*). To constrain such sources, we computed corresponding 95% upper limits on A_{gwb} at different values of α (Fig. 3). Other models are not well described by power laws, but their largest predicted signals are in or near the PTA band (*32*, *33*).

We have used the Fermi-LAT dataset to construct a gamma-ray PTA. This provides an independent method to search for signals detected with radio PTAs. Unlike the radio PTAs, this method is free from the effects of the IISM. Most of the pulsars are amenable to the TOA-based approach, and the resulting datasets are small compared with those of radio PTAs, enabling analysis alongside radio PTA data with little additional computational burden.

REFERENCES AND NOTES

1. A. Hewish, S. J. Bell, J. D. H. Pilkington, P. F. Scott, R. A. Collins, *Nature* **217**, 709–713 (1968).
2. C. M. Will, *Living Rev. Relativ.* **17**, 4 (2014).
3. B. P. Abbott *et al.*, *Ann. Phys.* **529**, 1600209 (2017).
4. A. Sesana, F. Haardt, P. Madau, M. Volonteri, *Astrophys. J.* **611**, 623–632 (2004).
5. S. Burke-Spolaor *et al.*, *Astron. Astrophys. Rev.* **27**, 5 (2019).
6. M. V. Sazhin, *Sov. Astron.* **22**, 36 (1978).
7. S. Detweiler, *Astrophys. J.* **234**, 1100 (1979).
8. R. W. Hellings, G. S. Downs, *Astrophys. J.* **265**, L39 (1983).
9. Z. Arzoumanian *et al.*, *Astrophys. J.* **905**, L34 (2020).
10. B. Goncharov *et al.*, *Astrophys. J. Lett.* **917**, L19 (2021).
11. S. Chen *et al.*, *Mon. Not. R. Astron. Soc.* **508**, 4970–4993 (2021).
12. J. Antoniadis *et al.*, *Mon. Not. R. Astron. Soc.* **510**, 4873–4887 (2022).
13. M. F. Alam *et al.*, *Astrophys. J. Suppl. Ser.* **252**, 4 (2021).
14. B. Goncharov *et al.*, *Mon. Not. R. Astron. Soc.* **502**, 478–493 (2021).
15. A. Melatos, B. Link, *Mon. Not. R. Astron. Soc.* **437**, 21–31 (2014).
16. A. Lyne, G. Hobbs, M. Kramer, I. Stairs, B. Stappers, *Science* **329**, 408–412 (2010).
17. J. M. Cordes, R. M. Shannon, D. R. Stinebring, *Astrophys. J.* **817**, 16 (2016).
18. J. Y. Donner *et al.*, *Astron. Astrophys.* **644**, A153 (2020).
19. R. M. Shannon, J. M. Cordes, *Mon. Not. R. Astron. Soc.* **464**, 2075–2089 (2017).
20. C. Tiburzi *et al.*, *Astron. Astrophys.* **647**, A84 (2021).
21. Materials and methods are available as supplementary materials.
22. W. B. Atwood *et al.*, *Astrophys. J.* **697**, 1071–1102 (2009).
23. M. Ajello *et al.*, *Astrophys. J. Suppl. Ser.* **256**, 12 (2021).
24. D. A. Smith *et al.*, *Astrophys. J.* **871**, 78 (2019).
25. M. Kerr, P. S. Ray, S. Johnston, R. M. Shannon, F. Camilo, *Astrophys. J.* **814**, 128 (2015).
26. L. Lentati *et al.*, *Mon. Not. R. Astron. Soc.* **437**, 3004–3023 (2014).
27. J. Ellis, M. Vallisneri, S. R. Taylor, P. T. Baker, ENTERPRISE: Enhanced Numerical Toolbox Enabling a Robust PulsaR Inference SuitE, *Zenodo* (2020); doi: 10.5281/zenodo.4059815.
28. X. Siemens, J. Ellis, F. Jenet, J. D. Romano, *Class. Quantum Gravity* **30**, 224015 (2013).
29. N. S. Pol *et al.*, *Astrophys. J. Lett.* **911**, L34 (2021).
30. W. Zhao, *Phys. Rev. D Part. Fields Gravit. Cosmol.* **83**, 104021 (2011).
31. T. Damour, A. Vilenkin, *Phys. Rev. D Part. Fields Gravit. Cosmol.* **71**, 063510 (2005).
32. Z. Arzoumanian *et al.*, *Phys. Rev. Lett.* **127**, 251302 (2021).
33. A. Khmelnitsky, V. Rubakov, *J. Cosmol. Astropart. Phys.* **2014**, 019 (2014).
34. M. Kerr, A. Parthasarathy, Software and data release for Fermi Pulsar Timing Array. *Zenodo* (2022).

ACKNOWLEDGMENTS

We dedicate this work to our recently deceased colleague, Jing Luo. We are grateful to the three insightful anonymous reviewers and to D. Champion for comments on an early draft. The Fermi-LAT Collaboration acknowledges support for LAT development, operation, and data analysis from NASA and the US Department of Energy (DOE) (United States); CEA/Irfu and IN2P3/CNRS (France); ASI and INFN (Italy); MEXT, KEK, and JAXA (Japan); and the K. A. Wallenberg Foundation, the Swedish Research Council, and the National Space Board (Sweden). Science analysis support in the operations phase from INAF (Italy) and CNES (France) is also gratefully acknowledged. The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities. Pulsar research at UBC is supported by an NSERC Discovery Grant and by CIFAR. Work at NRL is supported by NASA. **Funding:** This work was performed in part under DOE contract DE-AC02-76SF00515. M.K. is supported by NASA grant NNG21OB03A. E.C.F. and N.M. are supported by NASA under award 80GSFC21M0002. T.C. is supported by NASA through the NASA Hubble Fellowship Program grant HST-HF2-51453.001. K.C. is supported by a UBC Four Year Fellowship (6456). S.M.R. is a CIFAR Fellow and is supported by the NSF Physics Frontiers Center award 1430284. The work of M.A.S.C. and V.G. was supported by grants PGC2018-095161-B-I00 and CEX2020-001007-S, both funded by MCIN/AEI/10.13039/501100011033 and by ERDF. V.G. has been supported by Juan de la Cierva-Incorporación IJC2019-040315-I grants. G.Z. acknowledges financial support from the Slovenian Research Agency (grants P1-0031, I0-0033, and J1-1700). C.J.C. acknowledges support from the ERC under the European Union's Horizon 2020 research and innovation programme (grant agreement 715051; Spiders). S.J.S. holds an NRC Research Associateship award at NRL. **Author contributions:** M.Ke. conceived the project, implemented the gamma-ray analysis, and cowrote the manuscript. A.P. implemented the TOA-based analysis and cowrote the manuscript. D.A.S., P.S.R., M.P.R., and M.Kr. internally reviewed the manuscript. B.B., I.C., H.C., K.C., L.G., M.J.K., S.M.R., J.R., R.S., I.S., S.T., and G.T. contributed radio timing solutions. Other coauthors acquired Fermi-LAT data and reviewed and contributed to the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All raw data, processed data, reduced data, pulsar timing solutions, and software developed for this work are available on Zenodo (*34*).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abm3231
Fermi-LAT Collaboration Authors
Materials and Methods
Figs. S1 to S3
Tables S1 to S8
References (*35*–*92*)

10 September 2021; accepted 24 March 2022
Published online 7 April 2022
10.1126/science.abm3231

EXTINCTION

Avoiding ocean mass extinction from climate warming

Justin L. Penn[1,2]* and Curtis Deutsch[1,2]*

Global warming threatens marine biota with losses of unknown severity. Here, we quantify global and local extinction risks in the ocean across a range of climate futures on the basis of the ecophysiological limits of diverse animal species and calibration against the fossil record. With accelerating greenhouse gas emissions, species losses from warming and oxygen depletion alone become comparable to current direct human impacts within a century and culminate in a mass extinction rivaling those in Earth's past. Polar species are at highest risk of extinction, but local biological richness declines more in the tropics. Reversing greenhouse gas emissions trends would diminish extinction risks by more than 70%, preserving marine biodiversity accumulated over the past ~50 million years of evolutionary history.

Human activities are altering the global climate, physically transforming habitats, and overexploiting ecosystems of land and sea (*1*, *2*). As a result, rates of species extinction have risen above natural background levels (*3*, *4*). Documented extinctions are largely confined to land, where industrial human impacts began earlier and remain more pervasive despite rapid growth in commercial fishing, marine pollution, and transport (*5*). Climate change may eventually eclipse direct local human threats by causing widespread habitat loss through changes in the thermal and chemical conditions that reach even the most remote biomes on Earth, including the deep and open ocean (*6*, *7*). However, climate's impact on global biodiversity is challenging to observe, especially for undersampled marine environments (*8*), and common statistical models are difficult to validate, particularly as new climate conditions emerge (*9*–*11*).

The potential for substantial biodiversity loss is illustrated by the fossil record, where long-term diversification (*12*, *13*) is punctuated by episodic extinctions of varying intensity (Fig. 1A and fig. S1A). The most extreme events, the "Big 5" mass extinctions, coincided with global environmental changes, although mechanisms driving biodiversity collapse remain uncertain (*3*). In the largest such event, the end-Permian "Great Dying," loss of more than two-thirds of marine animal genera reduced biodiversity to near its minimum since animals first radiated (*12*–*14*). Similar environmental changes that occurred in the end-Permian, including rising temperatures and declining ocean O_2, productivity, and pH, are now also underway in the Anthropocene (fig. S2; *6*, *7*, *14*–*17*). Past mass extinctions provide the only empirical benchmarks for evaluating the severity and drivers of a "sixth mass extinction" (*3*, *18*).

Here, we project global and local extinction risks for marine animals (as a percentage of species lost) on the basis of habitat loss from climate change using an ecophysiological model that predicted the severity and latitude pattern of the end-Permian extinction (*17*, *19*). Multicentury climate and ocean conditions are simulated by a group of Earth system models that reproduce historical global warming trends (table S1; *6*, *20*) and project future climate changes under scenarios of high and low anthropogenic greenhouse gas emissions (Fig. 1C, inset; *6*, *7*, *21*). Global marine biodiversity is represented through a set of $>10^4$ simulated animal species types ("ecophysiotypes"; *17*), defined by thermal and hypoxia tolerance traits (table S2 and fig. S3) that structure the current biogeography of diverse taxa (*22*–*24*). Model species are assigned traits with frequencies found in a global species data compilation (*22*), allowing predicted patterns and climatic changes in biological richness to be compared with and tested against richness observations.

Ocean habitability requires conditions to remain within a species' ecophysiological tolerances. Local O_2 must meet temperature-dependent metabolic demands for growth and ecological activity, imposing covarying upper temperature and lower O_2 limits (*22*, *23*). A loss of aerobic habitat causes local species extirpation wherever ocean warming and O_2 loss drive the O_2 supply-to-demand ratio (Φ) below a species' critical threshold (Φ_{crit}) (*17*, *24*, *25*). Aerobic habitat losses of model species are comparable to those within the known geographic ranges of real species with corresponding traits (fig. S4). New habitat can also be gained at the cold edge of a species' range if temperatures there rise above its minimum tolerance (fig. S5; *25*, *26*).

Species are committed to global extinction if net habitat loss exceeds a critical fraction (V_{crit}) beyond which a viable population cannot be sustained even if disappearance is gradual (i.e., the extinction debt) (*27*). This uncertain extinction threshold is varied around a central value that is calibrated by the end-Permian extinction (fig. S6; *17*, *19*). We also considered the widest possible range of species capacities to colonize new habitat, denoted 0 and 100% colonization (Fig. 1B; *19*), using the median of these scenarios as our central extinction case.

As anthropogenic climate change accelerates, so too do projected species losses (fig. S7). The rate of these losses over time depends on the emissions scenario, which sets the pace of warming and ocean O_2 loss. By the latter half of this century, the divergent greenhouse gas emission scenarios will lead to markedly different climate trajectories and to growing disparities in the fraction of species lost globally (extinction) and locally (extirpation).

The eventual intensity of species losses for all emissions scenarios, climate models, and time periods is well predicted by the magnitude of global surface warming (Fig. 1, B and C). Total warming in turn is governed by cumulative emissions and varies across models with different climate sensitivities. For a given temperature change, extinction risk depends on the amount of O_2 lost (R^2 = 0.69 to 0.87, range across colonization scenarios), which also differs among models (fig. S8).

Under the low-emissions scenario, global temperature stops rising after ~1.9° ± 0.5°C of warming [intermodel mean, SD; (*6*)] by the end of the century, and species losses remain close to current commitments (Fig. 1B, inset, and C). Under the high-emissions scenario, surface air warming could reach ~4.9° ± 1.4°C by 2100 CE and ~10° to 18°C over the next three centuries (*6*), markedly elevating losses (Fig. 1, B and C). Significant losses are also expected in another high-emissions scenario (SSP3-7.0), which yields 8.2°C of warming by 2300 CE (5.7° to 11.8°C, 5 to 95% range) in a reduced-complexity model that does not simulate O_2 (*6*). Global extinction risk is higher if species cannot gain new habitat (Fig. 1B). Colonization of new regions at species' cold-edge range boundaries partially compensates for aerobic habitat loss (*25*) but eventually ceases to maintain habitat as warming intensifies.

We compared extinction risks from future climate change with those from current direct anthropogenic threats (fig. S9; *19*) on the basis of vulnerability assessments from the International Union for Conservation of Nature (IUCN; *28*). Future species losses from warming would become comparable to the sum of all anthropogenic stressors at the end of this century (Fig. 1B, inset, and C).

Under the high-emissions scenario, global extinction risks from continued warming eventually rival the severity of past mass extinctions in the fossil record and in paleosimulations

[1]School of Oceanography, University of Washington, Seattle, WA 98195, USA. [2]Department of Geosciences, Princeton University, Princeton, NJ 08544, USA.
*Corresponding author. Email: jpenn@princeton.edu (J.L.P.); cdeutsch@princeton.edu (C.D.)

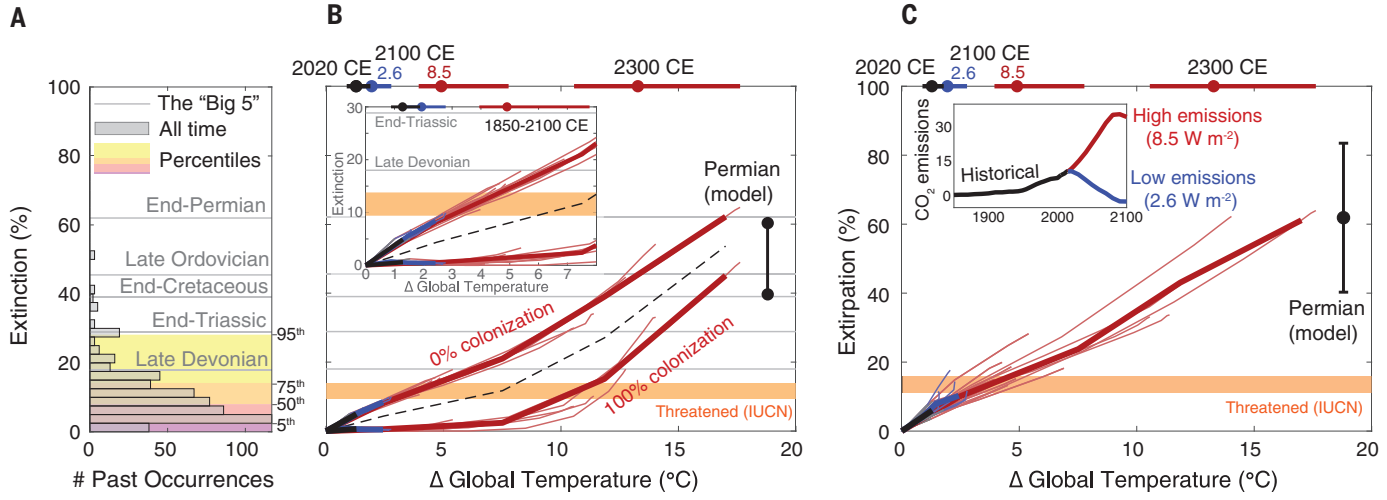


Fig. 1. Marine extinctions in the past and risk from climate warming. (**A**) Extinction intensity (percent losses) from the fossil record of marine animal genera over the past ~542 million years (*12*), including the "Big 5" mass extinctions (*35*) (fig. S1A). (**B** and **C**) Projected global extinction (B) and global mean extirpation risks (averaged over 1° latitude × longitude, 0 to 500 m) (C) rise with increases in annual mean global surface air temperature (thick lines are intermodel averages; thin lines are individual Earth system models) and are plotted under historical greenhouse gas emissions (petagrams of carbon per year) and divergent future scenarios [(C), inset], yielding radiative forcings of 2.6 W/m^2 [i.e., RCP/SSP1-2.6; (*6*, *19*, *21*)] versus 8.5 W/m^2 in 2100 [i.e., RCP/SSP5-8.5; (*6*, *19*, *21*)]. Inset in (B) zooms in to show end of the century changes. Percent colonization refers to the fraction of new habitat that a species can instantly disperse to and inhabit (thin dashed line is the median case). Extinction risks from current anthropogenic threats are estimated from IUCN vulnerability assessments (orange bars; tables S3 and S4; *19*). Simulated end-Permian extinction risks are shown across colonization scenarios in (B) and in (C) display the global mean and spatial SD of extirpation. Points and horizontal lines on the top axis show the average and range of warming across Earth system models evaluated here, respectively, for different time horizons and emissions scenarios.

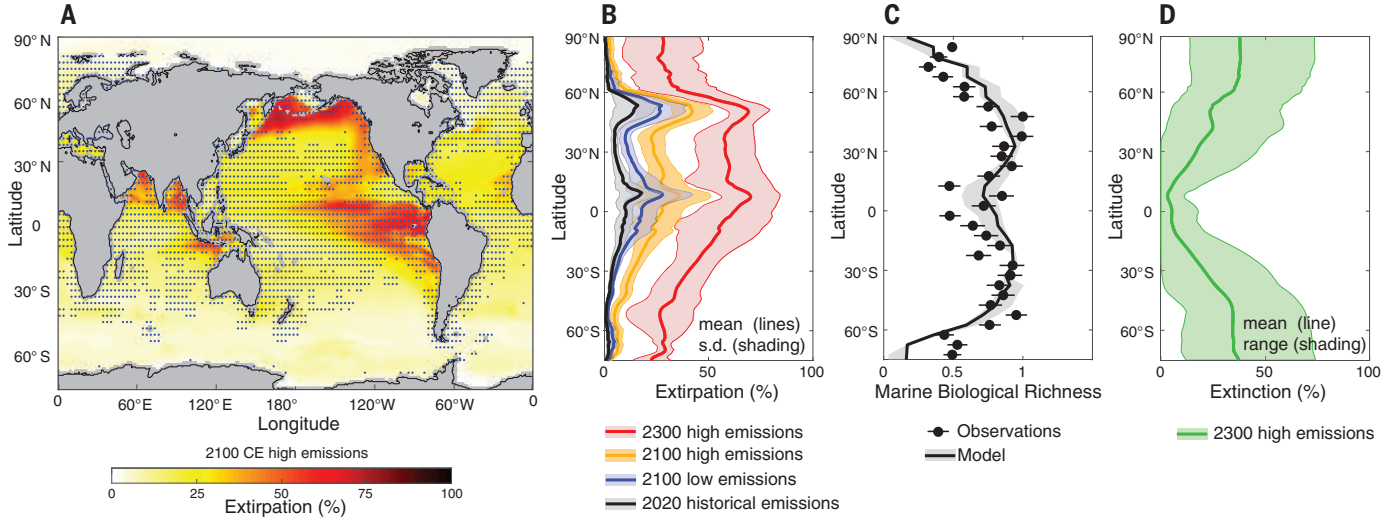


Fig. 2. Spatial variation in species losses and marine biological richness. (**A**) By 2100, regions of strong extirpations overlap past productive fisheries (blue points), where catch rates exceed the global median from 1950 to 2014 (*36*). (**B** to **D**) Patterns of extirpation risk (B), marine biological richness (C), and global extinction risk [(D); averaged across colonization scenarios] are shown versus latitude. Observed biological richness (number of species) estimated using rarefaction (*31*) is reproduced by the trait-based habitat model applied to climatological distributions of temperature and O_2 (*37*, *38*) across a range of maximum summation depths [(C) line is 500 m; shading is 0 and 5000 m; (*19*)]. Richness is normalized to the maximum observation. Extirpation and global extinction risks are averaged from 0 to 500 m, across Earth system models, and across longitude in (B) and (D).

(Fig. 1B). A mass extinction in this emissions scenario is projected across all potential values of the extinction threshold except in the unlikely case that the average species can maintain a viable population in <10% of their initial habitat volume (fig. S10; *19*). By contrast, limiting warming to 2°C would cut the severity of extinctions by >70%, avoiding a marine mass extinction across all extinction thresholds.

Climate-driven species losses vary widely among ocean biomes, with important implications for fisheries and the patterns of biological richness (Fig. 2). These spatial patterns also provide crucial observational tests of the underlying model.

Extirpation risk is greatest where climate anomalies are strong (fig. S2) or species are living close to their ecophysiological thresholds, typically equatorward near the warm and/or low-O_2 edges of their ranges (Fig. 2, A and B). Vulnerable regions include highly productive ecosystems where background O_2 is already low, such as the north Pacific, eastern boundary

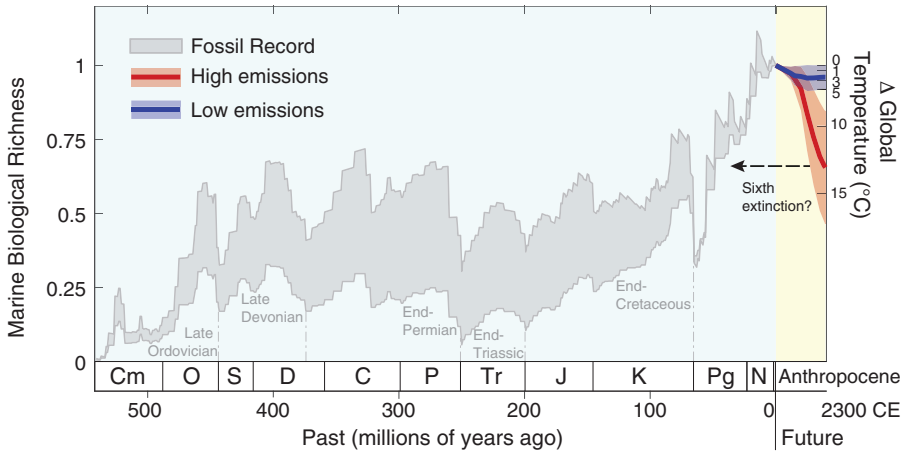


Fig. 3. Past and potential futures of marine biological richness. The number of marine animal genera is plotted over time from the fossil record relative to present and projected into the future on the basis of model extinction risks, averaged (lines) and varying (SD, shadings) across Earth system models and colonization scenarios. The right axis shows the change in global mean temperature for a given richness loss. Vertical dashed lines denote the "Big 5" mass extinctions. The lower fossil curve is based on Sepkoski's *Compendium of Fossil Marine Animal Genera* (*12*), and the upper curve presents these data with the secular trend found in the Paleobiology Database (https://paleobiodb.org) (fig. S12; *13*). Letters indicate major stratigraphic intervals plus the current Anthropocene. Time scale differs for past and future (see fig. S1B for both plotted on the same scale).

upwelling systems, and the tropical Indo-Pacific (Fig. 2A). These regions are also home to many of the world's most productive fisheries, which supply ~17% of humanity's dietary protein (*29*). The spatial pattern of extirpation is relatively stable over time, but its magnitude steadily rises with warming (Fig. 2B). Low extirpation intensities are predicted to be evident at low and northern midlatitudes already (Fig. 2B) and may underlie the previously documented range shifts observed there (*24*–*26*, *30*). Earth system models may underrepresent both small-scale spatial refugia as well as short-term heat waves and related extreme events. Although these phenomena will undoubtedly modulate local impacts on some species, they are unlikely to override the broader biotic outcome from the persistent large-scale and long-term climate trends presented here.

The latitudes of strong fractional species extirpation also overlap regions of peak biological richness (Fig. 2, B and C). The number of observed marine animal species increases from the poles toward the tropics, with a reduction near the equator (*31*–*33*). The model reproduces this pattern: As temperatures rise above species' minimum tolerances, richness increases from the poles to the tropics but decreases approaching the equator, where species reach their temperature-dependent hypoxia limits.

In contrast to local extinctions, global extinction risk is greater for polar species than for tropical ones, threatening higher-latitude richness where extirpation is weaker (Fig. 2, C and D). Species initially inhabiting the tropics can tolerate warm, low-O_2 waters, making them resilient to the climatic expansion of those conditions, especially for species with high colonization ability (fig. S11). By contrast, polar species occupy a disappearing climate niche and lack habitat refugia as the climate warms. This latitudinal extinction pattern has been detected in the fossil record of the end-Permian extinction (*17*, *34*), supporting the mechanistic basis for model projections.

The projected impact of accelerating climate change on marine biota is profound, driving extinction risk higher and marine biological richness lower than has been seen in Earth's history for the past tens of millions of years (Fig. 3 and fig. S1). Additional climate-related threats beyond warming and O_2 loss, including ocean acidification and declining primary productivity, have the potential to amplify these losses even further (*6*, *7*, *18*). However, it is not too late to enact the reductions in greenhouse gas emissions needed to avoid a major extinction event. The low-emissions scenario assumes that declines began around 2020 CE and continue thereafter (Fig. 1C, inset). Coordinated efforts to slow the local impacts of overfishing and marine pollution demonstrate the high intrinsic and economic value of preserving marine biodiversity and living resources. Realizing the fruits of these conservation efforts depends on mounting complementary societal responses to avert the greater threat posed by climate change.

REFERENCES AND NOTES

1. S. Díaz *et al.*, *Science* **366**, eaax3100 (2019).
2. S. L. Maxwell, R. A. Fuller, T. M. Brooks, J. E. M. Watson, *Nature* **536**, 143–145 (2016).
3. A. D. Barnosky *et al.*, *Nature* **471**, 51–57 (2011).
4. G. Ceballos *et al.*, *Sci. Adv.* **1**, e1400253 (2015).
5. D. J. McCauley *et al.*, *Science* **347**, 1255641 (2015).
6. Intergovernmental Panel on Climate Change (IPCC), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge Univ. Press, 2021).
7. L. Kwiatkowski *et al.*, *Biogeosciences* **17**, 3439–3470 (2020).
8. C. Mora, D. P. Tittensor, S. Adl, A. G. B. Simpson, B. Worm, *PLOS Biol.* **9**, e1001127 (2011).
9. W. W. L. Cheung *et al.*, *Fish Fish.* **10**, 235–251 (2009).
10. J. G. Molinos *et al.*, *Nat. Clim. Chang.* **6**, 83–88 (2016).
11. C. M. Beale, J. J. Lennon, A. Gimona, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 14908–14912 (2008).
12. R. A. Rohde, R. A. Muller, *Nature* **434**, 208–210 (2005).
13. J. Alroy *et al.*, *Science* **321**, 97–100 (2008).
14. J. L. Payne, M. E. Clapham, *Annu. Rev. Earth Planet. Sci.* **40**, 89–111 (2012).
15. H. Jurikova *et al.*, *Nat. Geosci.* **13**, 745–750 (2020).
16. S. E. Grasby, B. Beauchamp, J. Knies, *Geology* **44**, 779–782 (2016).
17. J. L. Penn, C. Deutsch, J. L. Payne, E. A. Sperling, *Science* **362**, eaat1327 (2018).
18. J. L. Payne, A. M. Bush, N. A. Heim, M. L. Knope, D. J. McCauley, *Science* **353**, 1284–1286 (2016).
19. Materials and methods are available as supplementary materials.
20. S. M. Papalexiou, C. R. Rajulapati, M. P. Clark, F. Lehner, *Earths Futur.* **8**, e2020EF001667 (2020).
21. B. C. O'Neill *et al.*, *Geosci. Model Dev.* **9**, 3461–3482 (2016).
22. C. Deutsch, J. L. Penn, B. Seibel, *Nature* **585**, 557–562 (2020).
23. C. Deutsch, A. Ferrel, B. Seibel, H. O. Pörtner, R. B. Huey, *Science* **348**, 1132–1135 (2015).
24. E. M. Howard *et al.*, *Sci. Adv.* **6**, eaay3188 (2020).
25. M. L. Pinsky, R. L. Selden, Z. J. Kitchel, *Ann. Rev. Mar. Sci.* **12**, 153–179 (2020).
26. J. M. Sunday, A. E. Bates, N. K. Dulvy, *Nat. Clim. Chang.* **2**, 686–690 (2012).
27. D. Tilman, R. M. May, C. L. Lehman, M. A. Nowak, *Nature* **371**, 65–66 (1994).
28. International Union for Conservation of Nature, "The IUCN Red List of Threatened Species, version 2019-3" (IUCN, 2019); https://www.iucnredlist.org/.
29. C. Costello *et al.*, *Nature* **588**, 95–100 (2020).
30. J. Lenoir *et al.*, *Nat. Ecol. Evol.* **4**, 1044–1059 (2020).
31. C. Chaudhary, H. Saeedi, M. J. Costello, *Trends Ecol. Evol.* **32**, 234–237 (2017).
32. D. P. Tittensor *et al.*, *Nature* **466**, 1098–1101 (2010).
33. M. Yasuhara *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **117**, 12891–12896 (2020).
34. C. J. Reddin, Á. T. Kocsis, W. Kiessling, *Paleobiology* **45**, 70–84 (2019).
35. S. M. Stanley, *Proc. Natl. Acad. Sci. U.S.A.* **113**, E6325–E6334 (2016).
36. D. Pauly, D. Zeller, "Sea Around Us concepts, design and data (2015); http://www.seaaroundus.org/.
37. R. A. Locarnini *et al.*, "World Ocean Atlas 2013, volume 1: Temperature" (NOAA, 2013); https://www.ncei.noaa.gov/data/oceans/woa/WOA13/DOC/woa13_vol1.pdf.
38. H. E. Garcia *et al.*, "World Ocean Atlas 2013, volume 3: Dissolved oxygen, apparent oxygen utilization, and oxygen saturation" (NOAA, 2014); https://www.ncei.noaa.gov/data/oceans/woa/WOA13/DOC/woa13_vol3.pdf.
39. Model output and code for: J. L. Penn, C. Deutsch, Avoiding ocean mass extinction from climate warming, Zenodo (2022); https://doi.org/10.5281/zenodo.5899743.

ACKNOWLEDGMENTS

We thank K. Armour for insightful discussions, H. Frenzel and G. Tsui for technical support, four anonymous reviewers for constructive feedback, and contributors to the following public data repositories: CMIP, IUCN, OBIS, PBDB, and the Sea Around Us. **Funding:** This work was supported by the National Science Foundation (grant OCE-1737282), the National Oceanic and Atmospheric Administration (grant NA18NOS4780167), the California SeaGrant and Ocean Protection Council, and the UW Program on Climate Change. **Author contributions:** J.L.P. performed simulations and analyses. Both authors designed the study and wrote the paper. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** See tables S1 and S5 for links to publicly available data used in this study. Model output and code are available at Zenodo (*39*).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abe9039
Materials and Methods
Figs. S1 to S12
Tables S1 to S5
References (*40*, *41*)
MDAR Reproducibility Checklist

2 October 2020; accepted 10 February 2022
10.1126/science.abe9039

ORGANIC CHEMISTRY

Scaffold hopping by net photochemical carbon deletion of azaarenes

Jisoo Woo[1], Alec H. Christian[2], Samantha A. Burgess[3], Yuan Jiang[3], Umar Faruk Mansoor[2], Mark D. Levin[1]*

Discovery chemists routinely identify purpose-tailored molecules through an iterative structural optimization approach, but the preparation of each successive candidate in a compound series can rarely be conducted in a manner matching their thought process. This is because many of the necessary chemical transformations required to modify compound cores in a straightforward fashion are not applicable in complex contexts. We report a method that addresses one facet of this problem by allowing chemists to hop directly between chemically distinct heteroaromatic scaffolds. Specifically, we show that selective photolysis of quinoline *N*-oxides with 390-nanometer light followed by acid-promoted rearrangement affords *N*-acylindoles while showing broad compatibility with medicinally relevant functionality. Applications to late-stage skeletal modification of compounds of pharmaceutical interest and more complex transformations involving serial single-atom changes are demonstrated.

The maturation of chemical synthesis has given rise to an era in which molecules can be exhaustively optimized to serve specific purposes under exceptional multidimensional constraints, enabling increasingly precise applications of these compounds. The intensity of this enterprise manifests most visibly in medicinal chemistry, where the simultaneous management of efficacy, specificity, absorption, and lifetime is accomplished through meticulous tailoring of promising candidate molecules (*1*, *2*). While these molecular optimizations establish structure-activity relationships by iterative modification of a series of parent compounds, the synthetic practice underlying these campaigns is rarely in line with its philosophical roots. Rather than converting a lead compound to the next candidate in a manner matching their underlying thought process, these campaigns instead largely rely on iterative resynthesis, because the reactions necessary to perform the envisioned direct conversion are often not applicable in complex settings (Fig. 1A) (*3*, *4*). This shortcoming is particularly conspicuous when conducting a "scaffold hop"—a common strategy that leverages computational estimates of three-dimensional molecular similarity (or in silico binding affinity to the target) to predict isofunctional structures with distinct cores (*5*, *6*). The logic of this strategy can be immediately appreciated by comparing members of a given class of pharmaceuticals, for example, the cholesterol-lowering therapeutics pitavastatin and fluvastatin or the anti-inflammatory drugs etoricoxib and celecoxib; this same logic is also clear when comparing compounds in a given development series, such as the dideazafolic acid antecedent to the chemotherapy agent pemetrexed (Fig. 1B) (*7*–*9*). Unfortunately, the execution of a predicted scaffold hop is among the most difficult of possible lead optimization strategies to perform directly. Unlike diversification strategies relying on robust, late-stage coupling reactions that can target some peripheral substructures of a lead molecule for rapid interrogation, molecular cores are far more challenging to examine in a similar fashion because of the often-distinct preparative methods for the relevant (hetero)cyclic frameworks. As such, chemists interested in examining a scaffold hop are typically required to resort instead to an effective reset of their synthetic campaign, beginning from scratch to traverse laterally in chemical space.

Accordingly, a pressing challenge and increasing recent area of focus for modern organic synthesis is the development of transformations that can address the molecular skeleton with precision and enable direct scaffold hops between distinct core substructures within a given class of compounds. Ideally, such transformations would enable control at the level of single-atom precision (Fig. 1C), with more sophisticated changes possible through iterative elementary skeletal modifications (*10*). Notable recent contributions from several groups have been reported in the context of saturated aliphatic heterocycles (*11*–*13*). Although the centrality of aromatic and heteroaromatic scaffolds in medicinal chemistry suggests a clear priority for similar azaarene interconversion strategies, the stability of aromatic systems poses a substantial challenge: The reactive species typically required to breach the core are often not compatible with densely functionalized druglike compounds and rarely promote precise, selective downstream chemistry (*14*–*16*). We report here a transformation that confronts this challenge, enabling a broadly applicable ring contraction of quinoline *N*-oxides and related azaarenes. Subsequent deacylation of the product *N*-acylindole allows this transformation to serve as a net carbon deletion (Fig. 1D).

This advance is built on the classical photochemistry of quinoline *N*-oxides (**1**), whose diverse rearrangement products were meticulously cataloged by Buchardt, Streith, Kaneko, and Albini (Fig. 2A) (*17*–*19*). Although *N*-acylindoles (**2**) and related hydration products have been observed arising from a limited set of substrates, more complex rearrangement products often predominate, including quinolones, 2- and 3-acylindoles (bearing noncleavable acyl groups), and 3-hydroxyquinolines. Product mixtures of these compounds are typically observed, and in many cases, seemingly minute perturbations to the substrate structure result in drastic changes to the product distribution (see figs. S20 to S28 and the associated discussion for a brief summary). Beyond this, the classical mercury (Hg) lamp irradiation conditions are incompatible with many complex quinolines of relevance to medicinal chemistry (an observation we have reproduced—see below). On the basis of prior mechanistic work on these and related photochemical transformations, we suspected that the undesired products were the result of secondary photoprocesses of the intermediate, formally antiaromatic 2,1- and/or 3,1-benzoxazepines (**3**) (*20*, *21*). We hypothesized that these two-photon by-products could be avoided using a milder, narrow-spectrum light source.

Indeed, we have found that the use of 390-nm light-emitting diodes (LEDs) in place of traditional mercury lamps substantially improves this classical photoreaction, turning an academic curiosity into a potential workhorse transformation with broad utility (*22*–*27*). This selective irradiation of quinoline *N*-oxides produces high yields of the corresponding 3,1-benzoxazepine in the photolysate, with subsequent in situ treatment with an acid catalyst promoting isomerization to the *N*-acylindoles in similarly high yield. As detailed below, this enables challenging indole and azaindole syntheses; facilitates late-stage, direct scaffold hopping of medicinal compounds; and serves as a productive springboard for further skeletal modification strategies.

We began our investigation with 2-methylquinoline *N*-oxide (**1a**) (Fig. 2B). Irradiation with a 390-nm LED in toluene at ambient temperature for 5 hours resulted in complete consumption of the quinoline *N*-oxide, affording the corresponding benzoxazepine **3a** in 91% nuclear magnetic resonance (NMR) yield, along with a minor quantity of the deoxygenation product **4a** (15:1 selectivity). Although **3a** was not isolable without substantial decomposition, its conversion could be monitored by LED-NMR, allowing measurement of a quantum

[1]Department of Chemistry, University of Chicago, Chicago, IL, USA. [2]Discovery Chemistry, Merck & Co., Inc., Boston, MA, USA. [3]Analytical Research and Development, Merck & Co., Inc., Boston, MA, USA.
*Corresponding author. Email: marklevin@uchicago.edu

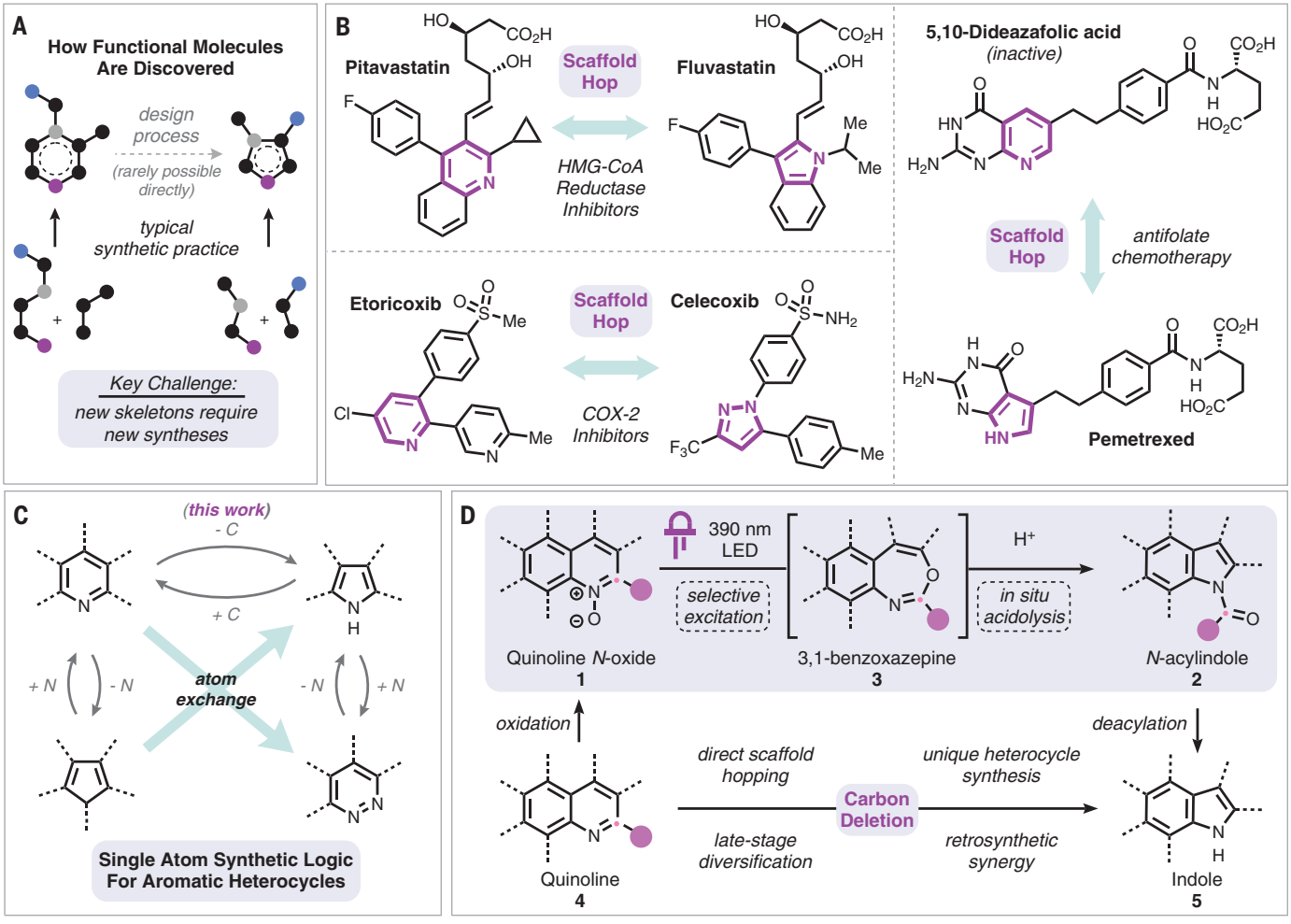


Fig. 1. Introduction to scaffold hopping and single-atom skeletal editing. (**A**) Schematic representation of the disconnect between design and synthesis in molecular optimization. (**B**) Selected examples of scaffold hopping in medicinal chemistry. HMG-CoA, 3-hydroxy-3-methylglutaryl coenzyme A; Me, methyl group; COX-2, cyclooxygenase-2. (**C**) Single-atom skeletal editing for heterocycle interconversion. (**D**) Carbon deletion of azaarenes delineated in this work.

yield, Φ, of 0.096 for conversion of the quinoline *N*-oxide (*28*). In contrast to **3a**, cyano-substituted benzoxazepine **3b** was found to be readily isolable and could similarly be prepared by irradiation with a 390-nm LED in 83% isolated yield. The increased stability of **3b** offered an opportunity to probe our light source hypothesis. As predicted, **3b** was found to be substantially more stable upon further irradiation by the LED than under mercury lamp irradiation (Fig. 2C). The benzoxazepine was returned in near-quantitative recovery after 6 hours in the former case, whereas nearly half of the material degraded to a mixture of products in the same period under the latter conditions. Examination of the absorption spectra of **1b** and **3b** reveals the origin of this drastic light source effect (Fig. 2D). Whereas the quinoline *N*-oxide has a relative absorption maximum, λ_{max}, at 386 nm, the benzoxazepine shows a substantial hypsochromic shift to a λ_{max} of 323 nm, such that the LED accomplishes selective irradiation of the starting material while the mercury lamp promotes photodegradation of the benzoxazepine.

Our interest in the photochemical behavior of quinoline *N*-oxides stemmed not from the benzoxazepine intermediates themselves but rather from the indole products presumed to arise from them. Classical studies had suggested that these indole products were the result of adventitious (or added) water generating acid, which in turn acted on the 3,1-benzoxazepine (*29*). This prompted us to examine the effect of exogenous acid additives. The crude photolysate consisting predominantly of methyl-substituted benzoxazepine **3a** reacted smoothly under the action of trifluoroacetic acid to afford acylindole **2a** in 78% yield relative to **1a**. To determine the mechanism by which the protonated benzoxazepine evolves to product, we conducted an ^{18}O labeling study, which showed substantial but incomplete maintenance of the isotopic label in the hydrolysis process. This result is most consistent with two concurrent pathways for benzoxazepine hydrolysis, although the potential for ^{18}O-water liberated in the *N*-protonation pathway to react further precludes a quantitative analysis of the partitioning between these pathways (Fig. 2E) (*30*).

Although direct photochemistry can often vary as a function of substrate structure, the efficacy and advantage of the 390-nm LED was found to be surprisingly general, both with respect to the substituent on the excised carbon and the residual indole substituents (Fig. 3). A wide range of quinoline *N*-oxides with varying substitution patterns were found to undergo facile photorearrangement to afford the corresponding benzoxazepine, and the subsequent acid-promoted rearrangement was likewise found to be generally applicable. For most substrates, trifluoroacetic acid was effective for this latter operation, although in cases bearing 3-substituents (**1m**, **1ac**, **1x**, and **1ag**) or electron-withdrawing groups (**1b** and **1s**) on the quinoline, the more acidic *para*-toluenesulfonic acid afforded higher yields. This protocol

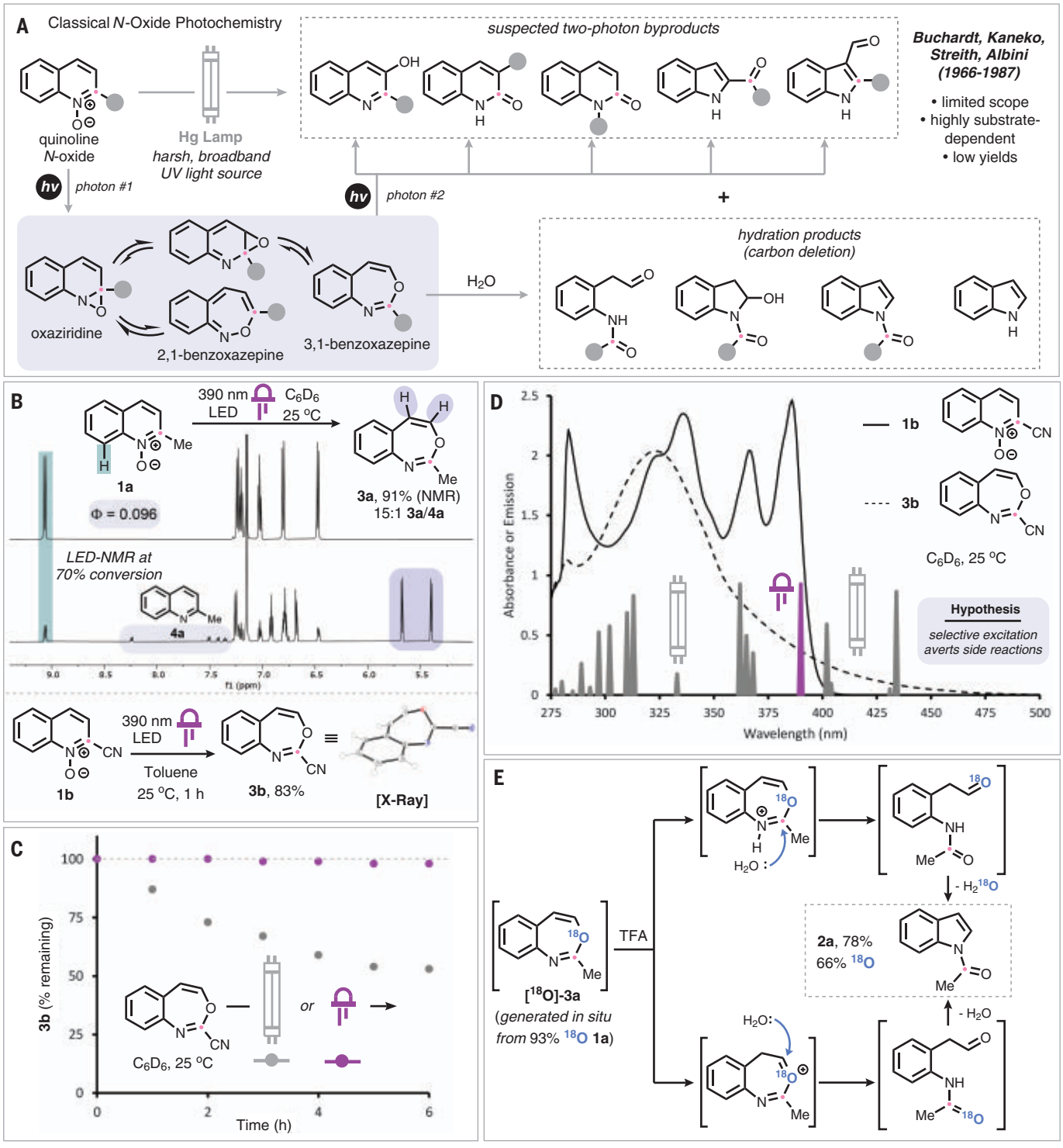


Fig. 2. Mechanistic basis for light source effects in quinoline *N*-oxide photochemistry and mechanism of acid-promoted benzoxazepine rearrangement to indole. (**A**) Summary of classical photochemical rearrangements promoted by broadband mercury lamp irradiation. (**B**) LED-NMR study of 390-nm LED photolysis of **1a** and isolation of 2-cyano-3,1-benzoxazepine (**3b**) from LED photolysate of quinoline *N*-oxide **1b**. ppm, parts per million. (**C**) Photostability study of **3b** under LED and Hg lamp irradiation. (**D**) Ultraviolet–visible absorption spectra of **1b** and **3b** with overlaid relative emission spectra of 390-nm Kessil lamp and 200-W medium-pressure Hg lamp (normalized to an emission intensity of 1 for λ_{max}). (**E**) Labeling study of trifluoroacetic acid (TFA)–promoted rearrangement of **3a**.

was also found to be amenable to a one-pot process, with initial formation of the *N*-oxide induced by hydrogen peroxide in dichloromethane followed by dilution of the crude reaction mixture with toluene (95:5) before irradiation affording the *N*-acylindole in only slightly diminished yields (69%) relative to the two-pot procedure with purified *N*-oxide (78%).

As noted above, this transformation was found to be robust to a wide variety of functionality commonly encountered in medicinal chemistry, including other heterocycles (**2d**, **2e**, **2f**,

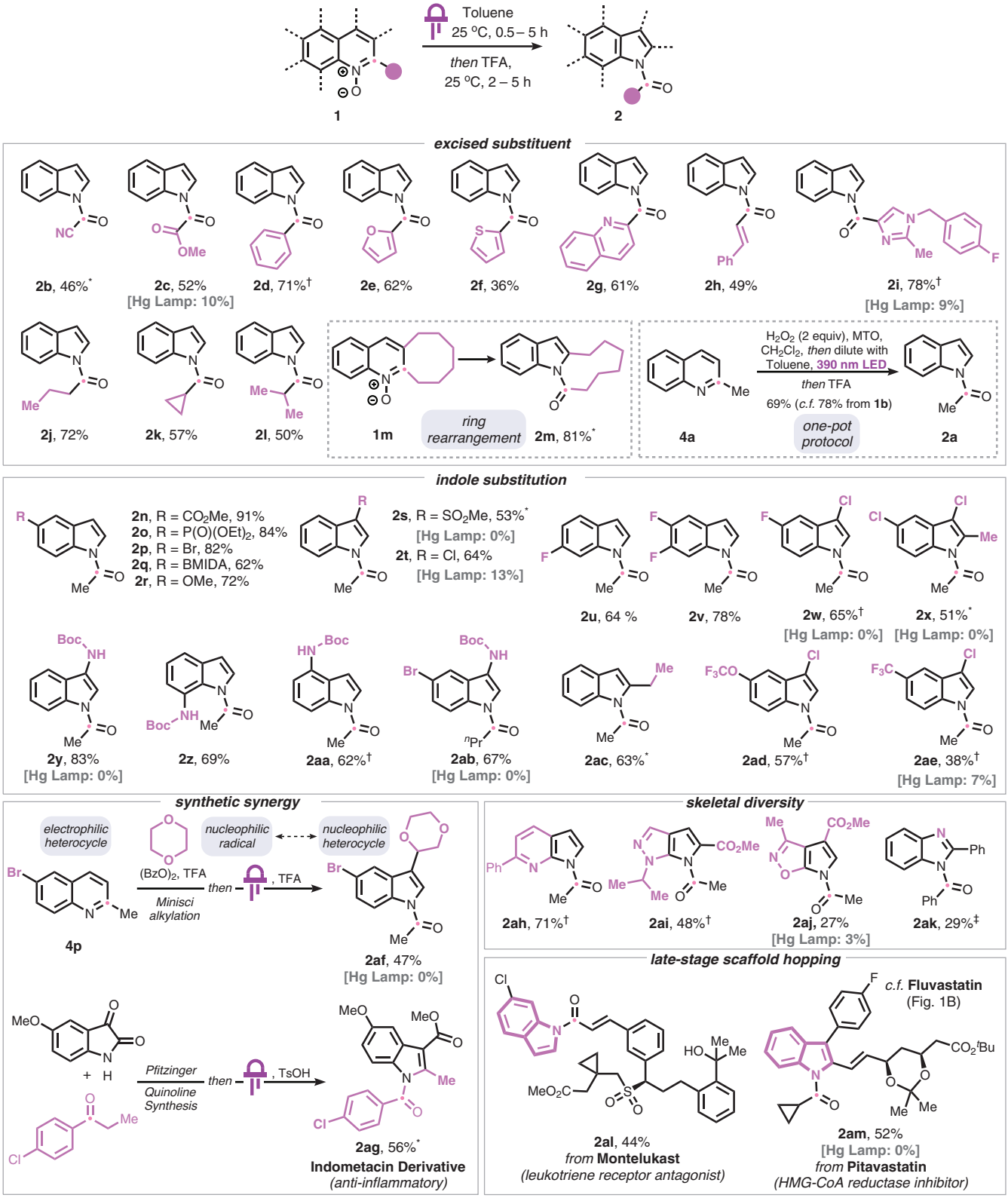


Fig. 3. Scope of the photochemical conversion of azaarene *N*-oxides into acylazoles. Conditions: **1** (0.3 mmol), toluene (0.06 M), 390-nm LED, 0.5 to 5 hours at 25°C, then TFA (0.3 mmol) for 2 to 5 hours at 25°C. Isolated yields are given. The asterisk symbol indicates where acidolysis was conducted with TsOH•H_2O instead of TFA. The single-dagger symbol indicates where acidolysis was conducted at 60°C. The double-dagger symbol indicates where, instead of acidolysis, photolysate was concentrated and refluxed in H_2O/MeOH (1:1). Hg lamp yields measured by proton nuclear magnetic resonance versus an internal standard. Ph, phenyl group; OMe, methoxy group; MTO, methyltrioxorhenium; BMIDA, *N*-methyliminodiacetyl boronate; Boc, butoxycarbonyl; ^{n}Pr, *n*-propyl; TsOH, *para*-toluenesulfonic acid; $(BzO)_2$, benzoyl peroxide; ^{t}Bu, *tert*-butyl.

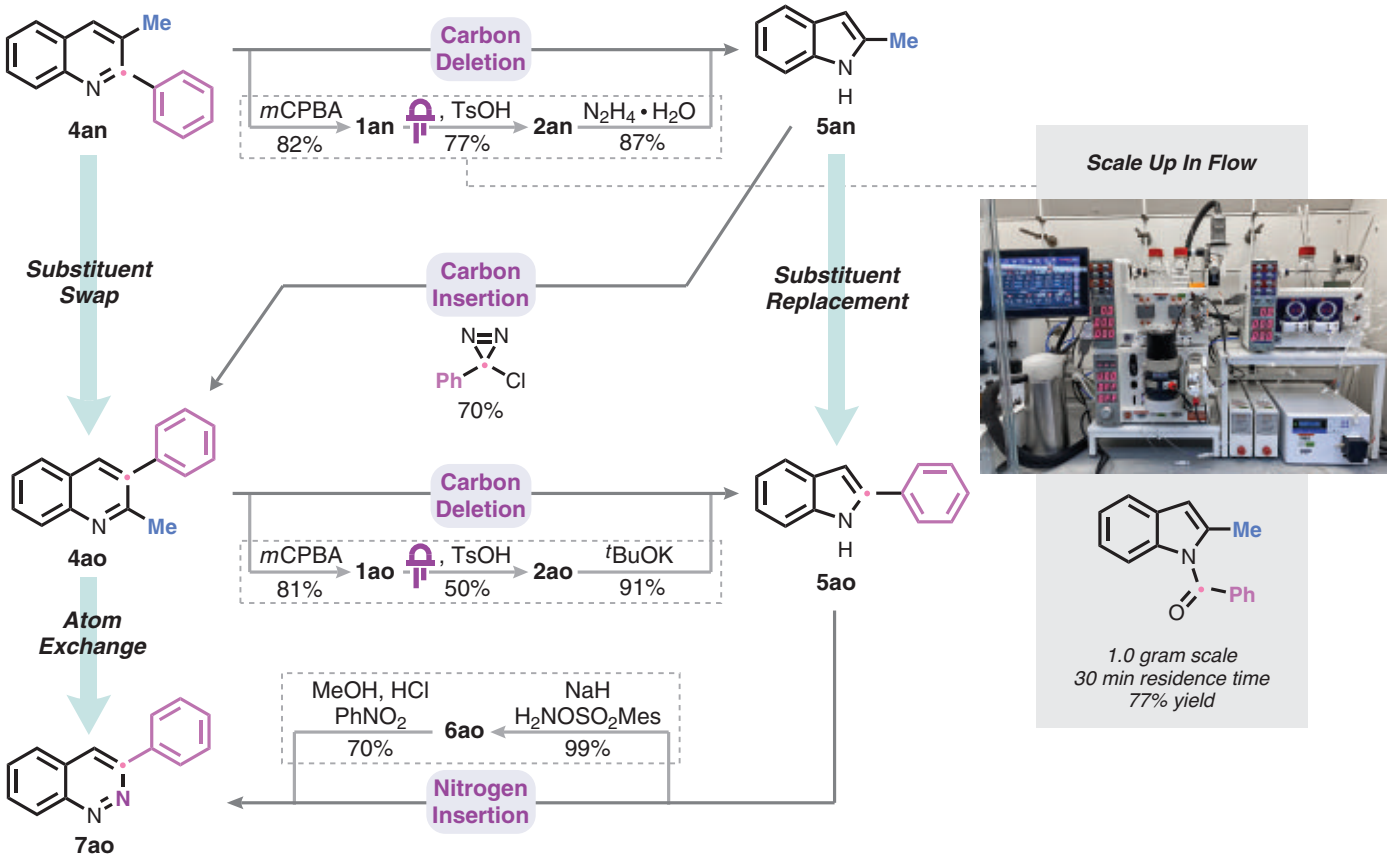


Fig. 4. Carbon deletion as a springboard for more complex scaffold hopping strategies and flow scale-up of photochemical rearrangement. Net substituent swaps of quinolines, substituent replacements of indoles, and C-to-N exchange of quinolines into cinnolines. See supplementary materials for detailed conditions. *m*CPBA, *meta*-chloroperoxybenzoic acid; tBuOK, potassium *tert*-butoxide; MeOH, methanol; HCl, hydrochloric acid; $PhNO_2$, nitrobenzene; NaH, sodium hydride; H_2NOSO_2Mes, *O*-mesitylsulfonyl hydroxylamine.

2g, and **2i**), polyhalogenation (**2v**, **2w**, and **2x**), carbamates (**2y**, **2z**, **2aa**, and **2ab**), phosphonates (**2o**), sulfones (**2s**), and boronic esters (**2q**). Notably, the scope demonstrated here was found to be a direct consequence of the milder light source. Of 12 quinolines examined, five gave detectable indole products at substantially diminished yield (3 to 13%) when a mercury lamp was used for the initial photolysis, with the remaining seven giving completely intractable mixtures. Additionally, complete solubility in toluene was not a requirement, with dissolution over the course of irradiation observed for a number of substrates (e.g., **1s**, **1q**, **1y**, **1af**, and **1aj**). In all cases, full conversion of the starting material was achieved with sufficient irradiation time. The most common by-product observed was deoxygenation of the *N*-oxide to afford the parent azaarene. Limitations of the photolysis were principally related to the 2-substituent (hydrogen, tertiary alkyls, and heteroatom substitution were not tolerated; see supplementary materials for further details). We also note that oxidatively sensitive functionality is not maintained in the initial *N*-oxidation (e.g., sulfides are converted to sulfones).

An interesting consequence of the switch between electron-poor quinoline and electron-rich indole heterocycles is the ability to interface their distinct reactivities and syntheses through carbon deletion. This is exemplified in the first instance by the preparation of indole **2af** through Minisci alkylation of the parent quinoline at the 4-position, resulting in the net 3-alkylation of the final, nucleophilic indole product with an nucleophilic radical—a challenging retrosynthetic strategy to realize via known methods (*31*–*33*). Quinoline **4ag** demonstrates the latter interplay of the two heterocyclic scaffolds, allowing the Pfitzinger quinoline synthesis to serve additionally as an indole synthesis (*34*). The product **2ag** is related to the anti-inflammatory medicine indomethacin (*35*).

Higher polyazaarenes were also found to be productive substrates, enabling the preparation of 7-azaindole, pyrrolopyrazole, pyrroloisoxazole, and benzimidazole scaffolds through net carbon deletion of the parent fused-ring azine. The 5,5-fused systems are highly challenging to prepare by traditional heterocycle syntheses and thus showcase a distinctive advantage of our approach. To further highlight the utility of this method, we demonstrated its capacity to modify complex medicinal compounds. Starting from montelukast (Singulair), a widely prescribed leukotriene inhibitor, the pendant chloroquinoline could be transformed into the corresponding acylindole **2al** (*36*). Finally, the direct scaffold hop from pitavastatin to its indole congener **2am** could be accomplished, creating a link in chemical space to fluvastatin via carbon deletion (*37*).

As noted above, linear combinations of distinct single-atom insertions and deletions offer exciting opportunities to devise more complex skeletal editing transformations. Figure 4 showcases the ways in which carbon deletion can be leveraged as a foundation for such strategies using a simple model system. Starting with quinoline **4an**, carbon deletion affords the indole **5an**, with the photorearrangement scalable up to 1 g in flow. Subsequent application of our previously reported C3-selective carbon insertion reaction gives the isomeric quinoline **4ao**, which has formally had its C2 and C3 substituents swapped relative to the starting **4an** (*38*). This quinoline can again be subjected to carbon deletion to afford indole **5ao**. Here, comparison to its predecessor **5an** reveals the effective replacement of the methyl

substituent with a phenyl. Finally, if indole **5ao** is subjected to nitrogen insertion through a precedented *N*-amination and oxidative aromatization sequence, cinnoline **7ao** can be accessed, now the formal C-to-N exchange product of starting quinoline **4ao** (*39*–*41*).

This work offers a broadly applicable, C2-selective, net carbon deletion of quinolines and related azaarenes through a ring contraction of the corresponding *N*-oxides. Avoiding deleterious overreaction through selective photoexcitation renders classical *N*-oxide photochemistry applicable to medicinal chemistry applications. This work further showcases the potential for direct scaffold hopping enabled by the carbon deletion transform, especially when used in combination with the growing library of single-atom skeletal edits.

REFERENCES AND NOTES

1. D. R. Owen *et al.*, *Science* **374**, 1586–1593 (2021).
2. I. J. P. de Esch, D. A. Erlanson, W. Jahnke, C. N. Johnson, L. Walsh, *J. Med. Chem.* **65**, 84–99 (2022).
3. D. C. Blakemore *et al.*, *Nat. Chem.* **10**, 383–394 (2018).
4. K. R. Campos *et al.*, *Science* **363**, eaat0805 (2019).
5. G. Schneider, W. Neidhart, T. Giller, G. Schmid, *Angew. Chem. Int. Ed.* **38**, 2894–2896 (1999).
6. Y. Hu, D. Stumpfe, J. Bajorath, *J. Med. Chem.* **60**, 1238–1246 (2017).
7. A. Endo, *Proc. Jpn. Acad. Ser. B Phys. Biol. Sci.* **86**, 484–493 (2010).
8. R. J. Flower, *Nat. Rev. Drug Discov.* **2**, 179–191 (2003).
9. E. C. Taylor, in *Successful Drug Discovery*, J. Fischer, D. P. Rotella, Eds. (Wiley 2015), pp. 157–180.
10. J. Jurczyk *et al.*, *Nat. Synth.* 10.1038/s44160-022-00052-1 (2022).
11. J. B. Roque, Y. Kuroda, L. T. Göttemann, R. Sarpong, *Nature* **564**, 244–248 (2018).
12. H. Lyu, I. Kevlishvili, X. Yu, P. Liu, G. Dong, *Science* **372**, 175–182 (2021).
13. S. H. Kennedy, B. D. Dherange, K. J. Berger, M. D. Levin, *Nature* **593**, 223–227 (2021).
14. E. Vitaku, D. T. Smith, J. T. Njardarson, *J. Med. Chem.* **57**, 10257–10274 (2014).
15. W. C. Wertjes, E. H. Southgate, D. Sarlah, *Chem. Soc. Rev.* **47**, 7996–8017 (2018).
16. X. Liu, C. Liu, X. Cheng, *Green Chem.* **23**, 3468–3473 (2021).
17. A. Albini, M. Alpegiani, *Chem. Rev.* **84**, 43–71 (1984).
18. G. G. Spence, E. C. Taylor, O. Buchardt, *Chem. Rev.* **70**, 231–265 (1970).
19. J. S. Poole, in *Heterocyclic N-Oxides*, O. V. Larionov, Ed., vol. 53 of *Topics in Heterocyclic Chemistry* (Springer, Cham, 2017), pp. 111–151.
20. E. E. Hurlow *et al.*, *J. Am. Chem. Soc.* **142**, 20717–20724 (2020).
21. C. Kaneko, A. Yamamoto, M. Hashiba, *Chem. Pharm. Bull. (Tokyo)* **27**, 946–952 (1979).
22. P. J. Sarver *et al.*, *Nat. Chem.* **12**, 459–467 (2020).
23. J. Jurczyk *et al.*, *Science* **373**, 1004–1012 (2021).
24. J. Ma *et al.*, *Science* **371**, 1338–1345 (2021).
25. P. Xu, P. López-Rojas, T. Ritter, *J. Am. Chem. Soc.* **143**, 5349–5354 (2021).
26. K. Targos, O. P. Williams, Z. K. Wickens, *J. Am. Chem. Soc.* **143**, 4125–4132 (2021).
27. G. Occhialini, V. Palani, A. E. Wendlandt, *J. Am. Chem. Soc.* **144**, 145–152 (2022).
28. Y. Ji, D. A. DiRocco, C. M. Hong, M. K. Wismer, M. Reibarkh, *Org. Lett.* **20**, 2156–2159 (2018).
29. J. Kurita, K. Iwata, T. Tsuchiya, *Chem. Pharm. Bull. (Tokyo)* **35**, 3166–3174 (1987).
30. M. L. Bender, R. D. Ginger, K. C. Kemp, *J. Am. Chem. Soc.* **76**, 3350–3351 (1954).
31. F. Minisci, E. Vismara, F. Fontana, *J. Org. Chem.* **54**, 5224–5227 (1989).
32. B. Giese, *Angew. Chem. Int. Ed.* **22**, 753–764 (1983).
33. D. Ma, B. S. Martin, K. S. Gallagher, T. Saito, M. Dai, *J. Am. Chem. Soc.* **143**, 16383–16387 (2021).
34. W. Pfitzinger, *J. Prakt. Chem.* **38**, 582–584 (1888).
35. F. D. Hart, P. L. Boardman, *BMJ* **2**, 965–970 (1963).
36. R. N. Young, in *Case Studies in Modern Drug Discovery and Development*, X. Huang, R. G. Aslania, Eds. (Wiley, 2012), pp. 154–195.
37. Y. Zhang *et al.*, *Org. Biomol. Chem.* **17**, 4970–4974 (2019).
38. B. D. Dherange, P. Q. Kelly, J. P. Liles, M. S. Sigman, M. D. Levin, *J. Am. Chem. Soc.* **143**, 11337–11344 (2021).
39. M. Somei, Y. Kurizuka, *Chem. Lett.* **8**, 127–128 (1979).
40. F.-L. Haut *et al.*, *J. Am. Chem. Soc.* **143**, 9002–9008 (2021).
41. S. Liu, X. Cheng, *Nat. Commun.* **13**, 425 (2022).

ACKNOWLEDGMENTS

We thank D. Nagib (Ohio State University) and C. Hong (Merck & Co., Inc., Kenilworth, NJ, USA) for helpful discussions. We thank J. Andrews (Hanovia/Colight) for providing spectral output data for the mercury lamp. We also thank A. Filatov, S. Whitmeyer, P. Kelly, and Y.-S. Chen for assistance with SXRD acquisition at beamline 15-ID-B,C,D. **Funding:** M.D.L. thanks the Packard Foundation and National Institutes of Health (R35 GM142768) for funding. X-ray diffraction data reported here were collected at ChemMatCARS Sector 15, which is supported by the NSF under grant NSF/CHE-1834750, and used resources of the Advanced Photon Source, a US DOE Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under contract DE-AC02-06CH11357. **Author contributions:** J.W. and M.D.L. conceived of the work. J.W. conducted synthetic experiments including purification and characterization. J.W., M.D.L., S.A.B., and A.H.C. designed the experiments reported in this work. S.A.B. conducted LED-NMR and quantum yield measurements. J.W., A.H.C., and Y.J. conducted the ^{18}O labeling experiment. A.H.C. and U.F.M. conducted the flow experiments. All authors prepared the manuscript. A.H.C. and M.D.L. directed the research. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** X-ray dataset for compound **3b** is freely available at the Cambridge Crystallographic Data Centre under deposition number 2159035.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abo4282
Materials and Methods
Figs. S1 to S49
Tables S1 to S4
NMR Spectra
References (*42*–*99*)

3 February 2022; accepted 25 March 2022
10.1126/science.abo4282

ORGANIC CHEMISTRY

Accelerating reaction generality and mechanistic insight through additive mapping

Cesar N. Prieto Kullmer[1]†, Jacob A. Kautzky[1]†, Shane W. Krska[2], Timothy Nowak[3], Spencer D. Dreher[2]*, David W. C. MacMillan[1]*

Reaction generality is crucial in determining the overall impact and usefulness of synthetic methods. Typical generalization protocols require a priori mechanistic understanding and suffer when applied to complex, less understood systems. We developed an additive mapping approach that rapidly expands the utility of synthetic methods while generating concurrent mechanistic insight. Validation of this approach on the metallaphotoredox decarboxylative arylation resulted in the discovery of a phthalimide ligand additive that overcomes many lingering limitations of this reaction and has important mechanistic implications for nickel-catalyzed cross-couplings.

Over the past century, organic chemists have invented a substantial number of catalytic bond–forming reactions (Fig. 1A). Many of these innovative transformations enable streamlined access to high-value molecular motifs. However, despite the vast and growing body of known chemical reactions, only a select few are routinely used by organic chemists and researchers in adjacent fields. Among these are olefin metathesis, the Suzuki coupling, and the Buchwald-Hartwig coupling (*1*–*3*).

The few transformations that have made the leap from invention to mainstay reaction share a key feature: They are high yielding and robust and capable of readily accommodating a wide range of substrate functionality and complexity. It is generally accepted that the elusive attributes of substrate generality and reaction efficiency cannot be anticipated, and it typically takes years of rigorous study to fully optimize the scope and yield of a challenging reaction (*1*–*7*). As outlined in Fig. 1, traditional reaction generalization is an iterative process that begins with careful mechanistic investigation. The insights gained in this exercise may suggest rational modifications that can lead to incrementally improved performance through elaborate catalyst optimization (*8*–*14*). Unfortunately, this approach becomes problematic when applied to inherently complex catalytic systems, in which mechanistic insights into underlying issues are not straightforwardly achieved or leveraged (hereafter referred to as complex reactions). As a result, the chemical literature is replete with promising but underutilized reactions that have yet to realize their

[1]Merck Center for Catalysis at Princeton University, Princeton, NJ 08544, USA. [2]Department of Process and Analytical Chemistry, MRL, Merck & Co., Inc., Rahway, NJ 07065, USA. [3]Department of Discovery Chemistry, MRL, Merck & Co., Inc., Kenilworth, NJ 07033, USA.
***Corresponding author. Email: dmacmill@princeton.edu (D.W.C.M.); spencer_dreher@merck.com (S.D.D.)**
†These authors contributed equally to this work.

full potential because of their mechanistic ambiguity.

In response to this challenge, we sought to develop an approach amenable to the generalization of complex reactions that also yields mechanistic insight that is inaccessible otherwise. Inspiration was drawn from the medicinal chemistry practice of phenotypic screening (*15*–*19*). A phenotypic screen involves the application of libraries of chemically diverse compounds to entire biological systems of interest while searching for desirable changes in the observable traits (the phenotype). This approach has the potential to simultaneously uncover and modulate hitherto unrecognized biomolecular mechanisms. It has proven invaluable in developing powerful therapeutic approaches that leverage unappreciated mechanisms of action and improved indications recalcitrant to traditional drug discovery techniques (*20*–*23*). We reasoned that complex reactions can be likened to complex biological systems, so there would be enormous benefit to applying this concept to the study and improvement of the former.

Unexpected improvements in reaction efficiency achieved with additives is a well-documented but often overlooked phenomenon within organic chemistry. Examples include the lithium chloride effect in Stille couplings (*24*) and, more recently, work conducted by the Watson group (*25*), the Dong group (*26*), and our own groups (*27*, *28*). Discovered additives can often be rationalized ex post facto but are nearly impossible to predict a priori. More importantly, knowledge of additive effects can be leveraged to extract important information about the mechanism of the reaction itself. Rapid evaluation of an additive library might therefore result in similar unforeseen mechanistic modifications that benefit the reaction and lead to important mechanistic insights.

Our envisioned strategy is outlined in Fig. 1 (bottom). We selected a challenging, complex reaction of limited substrate scope and began evaluating additives in a systematic fashion. Using high-throughput experimentation (HTE) methods, we were able to identify privileged motifs by mapping the evaluated additive space onto the resulting yield. Identified hits were then evaluated through structure-activity relationship (SAR) and mechanistic studies with the goal of identifying the optimal additive and gaining broader chemical insights. Rationalization of the additive effect should generate new, nonobvious mechanistic information useful to organic methodology at large.

We sought to apply this approach to the construction of $C(sp^2)$–$C(sp^3)$ bonds in an effort to forward the long-sought-after goal to "escape from flatland" within contemporary drug discovery (*29*, *30*). Several elegant methods have been developed to address this gap,

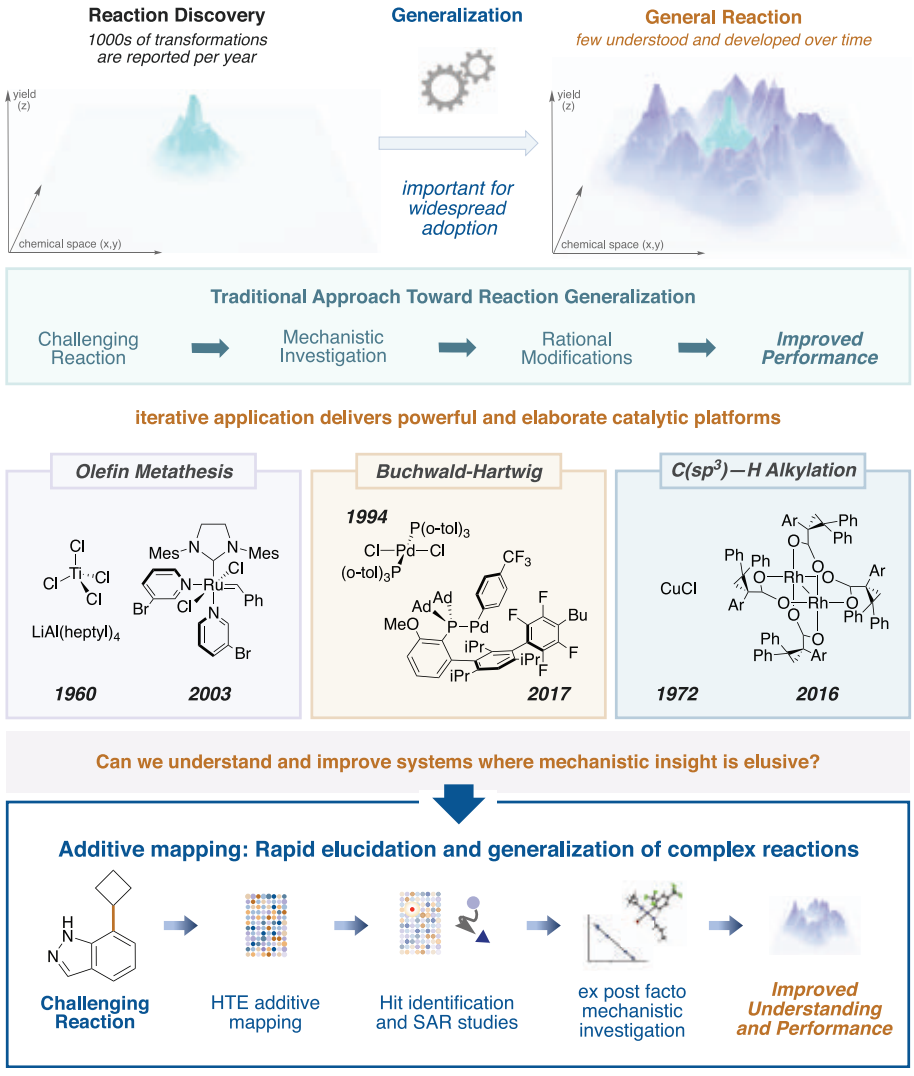


Fig. 1. High-throughput additive mapping to understand and improve organic methods. Shown are the traditional approach and select successes, as well as the proposed HTE additive mapping approach.

but none has the requisite starting material availability to enable broad access to chemical space (*31*–*37*). The metallaphotoredox decarboxylative arylation reaction initially published by the Doyle and MacMillan groups (*38*) has the putative advantage of broad commercial availability of both substrates (carboxylic acid and aryl halide) but has not gained widespread traction because of shortcomings in reaction generality. Specifically, the decarboxylative arylation is not amenable to (i) coordinating substrates, (ii) aryl bromides prone to protodehalogenation, (iii) challenging oxidative additions, or (iv) nonactivated carboxylic acids—namely, those that result in unstabilized radicals upon oxidative decarboxylation (*39*). Major efforts by our group and others (*40*) using traditional optimization approaches have failed to either substantially expand the scope of this reaction or to yield any helpful mechanistic leads. Therefore, our additive mapping approach could be well equipped to address this challenge.

To begin, we focused on an additive library of 721 diverse organic molecules amenable to HTE evaluation (Fig. 2B; also see the supplementary materials for further discussion). In principle, this approach should also be compatible with other classes of additives (e.g., salts, ligands, and metals). Concomitantly, we also selected several challenging coupling partners, including substrates with coordinating basic nitrogens, nonactivated carboxylic acids, and aryl halides that lead to sizeable quantities of Minisci and protodehalogenation side products. These reactions were performed in a nanomole-scale photoredox setup in which performance, as determined by reaction yield, was evaluated in the presence of each additive.

Unsurprisingly, most additives, particularly those containing groups such as heterocycles,

anilines, and phenols, led to attenuated reaction efficiency (*41*) (Fig. 2B). However, certain five-membered cyclic imides and hydantoins were found to give a strong boost in overall yield, sometimes up to fivefold, as well as a sharp decrease in protodehalogenation. This result was unexpected given that coordinating functionality is traditionally a powerful reaction poison caused by catalyst chelation.

Further evaluation of these unexpected hits was done using SAR studies. A total of 48 substrate combinations (three acids against 16 aryl bromides) were selected and evaluated on a nanoscale against 64 commercially available imide and hydantoin additives (Fig. 2C, left, and figs. S10 and S11). Substrate-dependent improvements were seen for a range of additives, but phthalimide proved to be the most broadly applicable. It generally gave the largest yield improvement, was commercially available, and was easily removed during workup. Methylation of the nitrogen, installation of electron-deficient substituents on the aromatic ring, or alteration of the five-membered ring size impaired or completely ablated the observed improvement. The steric effects of the imide additive, however, proved unimportant, with even tetramethylsuccinimide performing well.

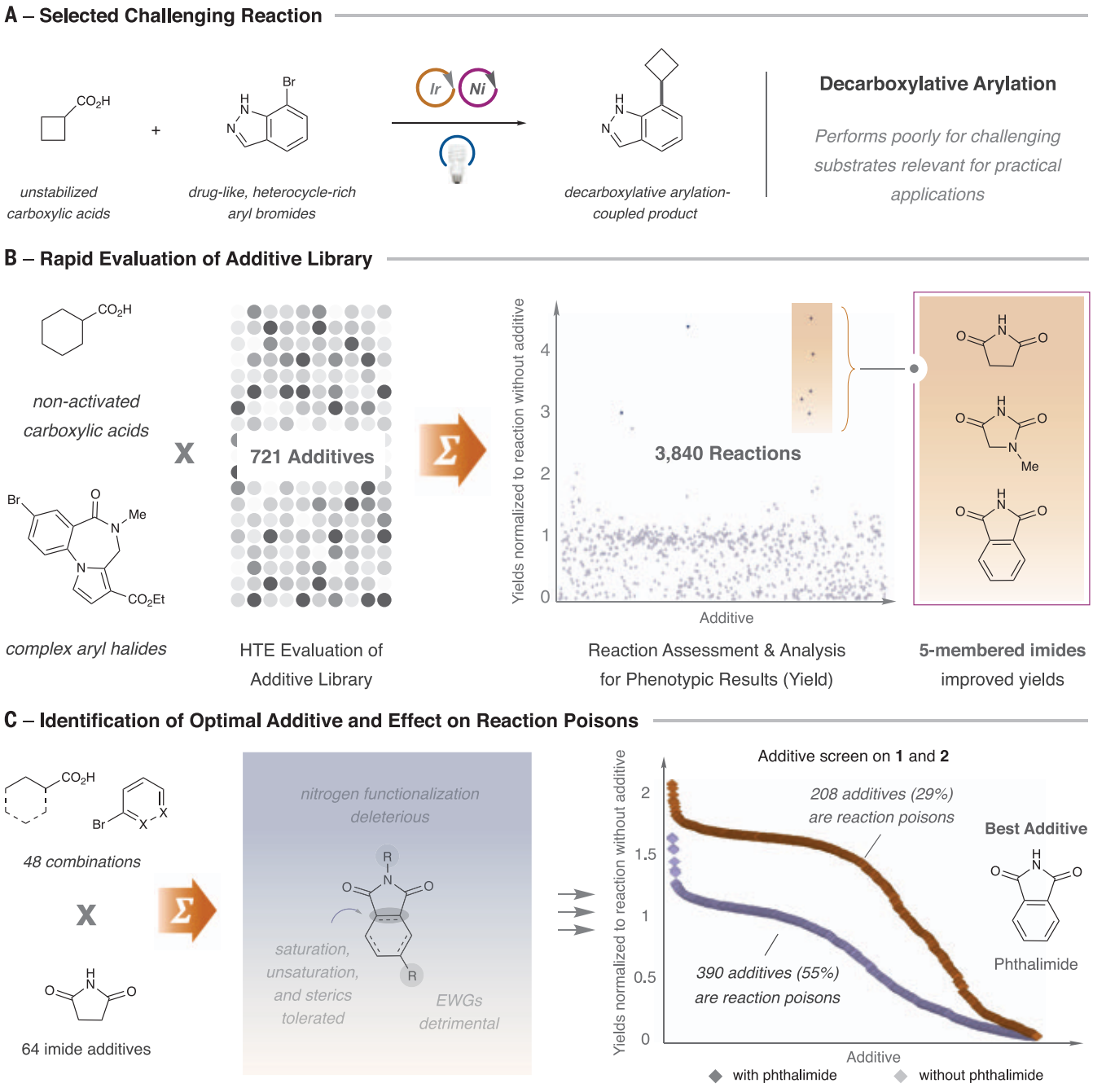


Fig. 2. Additive mapping applied to the decarboxylative reaction. (**A**) Metallaphotoredox decarboxylative arylation. (**B**) High-throughput additive screening on challenging coupling partners revealing the imide effect. (**C**) Left: SAR studies to identify the ideal five-membered imide. A set of 64 imides was evaluated in 48 couplings to reveal the best imide. Right: Phthalimide can counteract the poisoning effect of problematic functionalities.

To gauge improvements in the functional group tolerance, compound X2 and cyclohexanoic acid were resubjected to the earlier additive screen in the presence of phthalimide, and overall reaction performance was substantially better (Fig. 2C, right, and fig. S13). A range of compounds that previously served as reaction poisons (e.g., 1,3-dicarbonyls and benzoic acids) were now well tolerated. Furthermore, examining the data in aggregate showed that the overall average yield nearly doubled in the presence of phthalimide, and the overall number of reaction poisons (defined as decreasing the yield by >33%) fell from 390 to 208.

With these exciting results in hand, we sought to benchmark the reaction improvement in a pharmaceutically relevant context

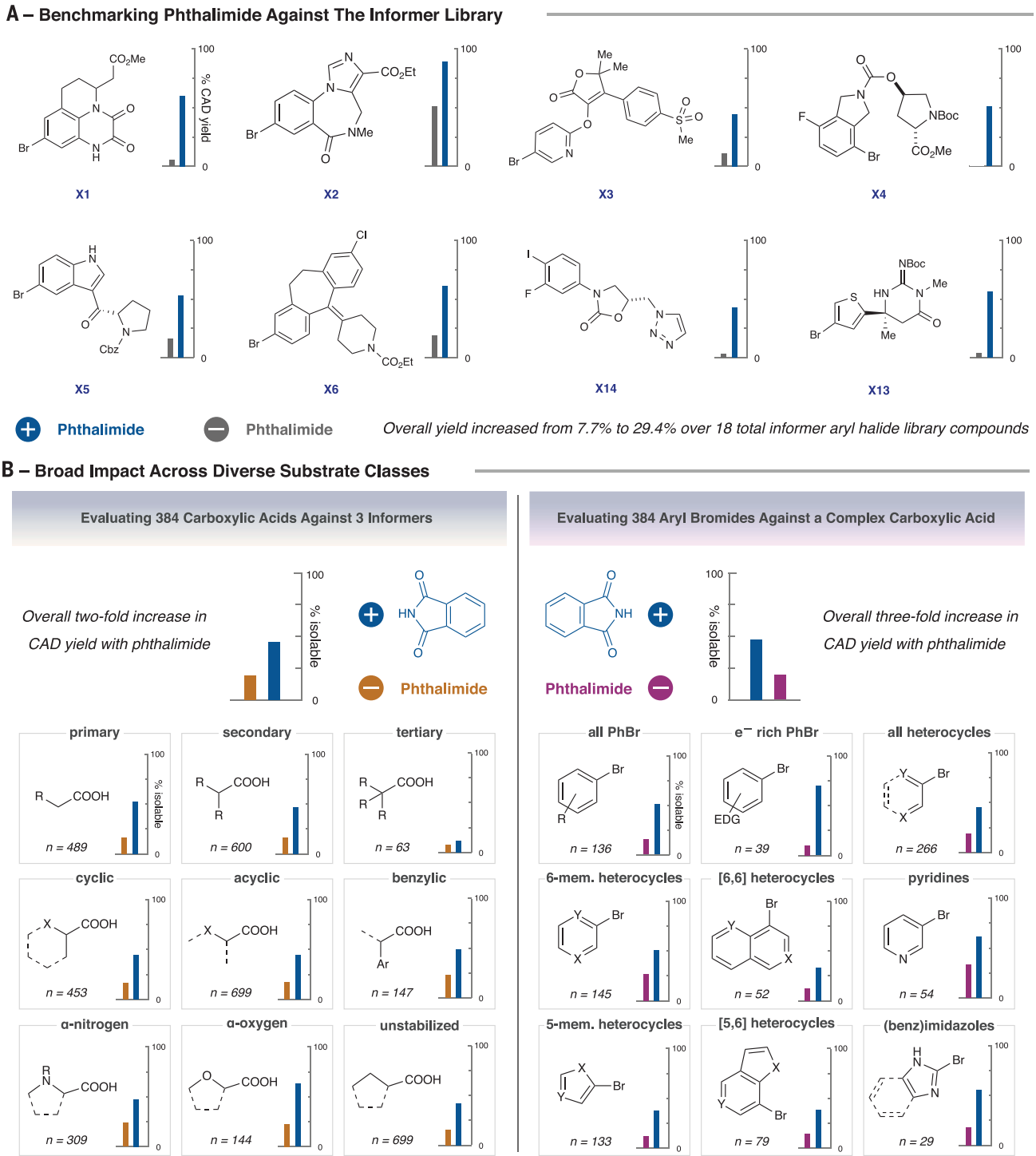


Fig. 3. Scope of the decarboxylative arylation. (**A**) Scope of the Aryl Halide Informer Library investigated against tetrahydropyran-4-carboxylic acid (% CAD yield reported). (**B**) Nanoscale high-throughput scope evaluation (% isolable reported). A product is deemed isolable if the CAD yield is ≥10%. *n*, number of evaluated compounds in this category; EDG, electron-donating group.

against the Aryl Halide Informer Library (*42*). This library is prototypical of the type of complex drug-like compounds seen in medicinal chemistry and contains a range of (hetero)aryl halides deemed to be inherently challenging for metal-catalyzed cross-couplings. When evaluating the library against a nonactivated acid, we found that phthalimide had an outstanding impact. A sizeable 11 of the 18 aryl halides showed major performance improvements, and a tripling of overall average yield was observed, from 7.7 to 29.4% (Fig. 3A and fig. S14). Failures included compounds with free carboxylic acids (X7 and X9) and aryl chlorides (X16 to X18), substrates that lie outside the scope of this transformation. The results position the decarboxylative arylation among the most successful couplings evaluated against the library.

Encouraged by the informer results, we then set our sights on evaluating the new scope of the reaction using the phthalimide additive. A highly diverse array of 384 small, medicinally relevant aryl bromides were evaluated against a complex, nonactivated isonipecotic acid derivative on a nanomole scale (see the supplementary materials for details). To gauge the synthetic utility of the phthalimide additive, we used charged aerosol detection (CAD) to determine the yield of each reaction (*43*). A 10% reaction yield was selected as threshold for the potential isolability of products by mass-directed microisolation that was based on prior work (*44*, *45*).

With phthalimide, the number of compounds above our typical threshold for isolation more than doubled, from 70 to 187 (Fig. 3B, right, and fig. S17). Substantial improvements were observed for a wide variety of bromides, including six-membered ring systems (aryl bromides, pyridines, and pyrimidines), five-membered heterocycles (pyrazoles, imidazoles, and thiazoles), and [6,5]- and [6,6]-heterobicycles (indoles, aza-indoles, benzimidazoles, and quinolines). Furthermore, phthalimide improved the functional group compatibility of the reaction, allowing for the presence of polar moieties such as 1,2-diols, phenols, and aminopyridines.

We next examined the scope of 384 chemically diverse, relevant carboxylic acids against the structurally complex informers X1, X2, and X13 on a nanomole scale (1152 total combinations). The phthalimide additive produced important improvements in a range of primary carboxylic acids including alpha ether, alpha thioether, benzylic, and, most notably, nonactivated acids, the latter representing a crucial advancement in overall reaction generality because they were previously only narrowly tolerated (Fig. 3B, left, and figs. S21 to S23). Further improvements were likewise seen for a range of cyclic carboxylic acids, including both activated and nonactivated four-, five-, six-, and seven-membered rings. A boost in yield was also seen for a range of acyclic carboxylic acids, and 19 of 20 protected amino acids were cross-coupled successfully in the presence of phthalimide, including both potential sites in aspartic acid and glutamic acid. Finally, the overall functional group compatibility of the reaction appeared far more robust in the presence of phthalimide. Substrates bearing a range of functionalities, including phenols, aldehydes, aryl chlorides, and β-alcohols, performed notably better with the additive. In aggregate, the number of reactions delivering products in >10% yields increased from 212 to 516, and the overall reaction CAD yield across the set increased more than twofold. The improvements in both aryl bromide and acid scope observed in this study should, in a realistic setting, have a large impact on the generality of the decarboxylative arylation in synthesis at large.

With these marked improvements in hand, we next sought to leverage the discovered effect of phthalimide to generate a fundamental mechanistic understanding useful for future Ni-metallaphotoredox developments. We first focused on elucidating how phthalimide turns nonactivated acids into competent coupling partners. Evaluating the reaction progress of a model system (figs. S31 and S32; substrate pair yields were 23% without phthalimide and 81% with phthalimide), we observed full consumption of the aryl bromide and substantial protodehalogenation in the absence of phthalimide. Conversely, suppression of protodehalogenation and predominant formation of the desired cross-coupled product was observed in the presence of phthalimide. We hypothesized that phthalimide prevents the decomposition of the intermediate oxidative addition complex (OAC), the most likely source of protodehalogenation.

To investigate the possibility of a stabilizing interaction between the putative Ni-aryl complex and phthalimide in the reaction, we sought to emulate the formation of such a complex under reaction-relevant conditions. The Ni(II) precatalyst was reduced by dropwise addition of $(Cp^*)_2Co$ in the presence of carboxylic acid, 2-tert-butyl-1,1,3,3-tetramethylguanidine (BTMG), and aryl bromide (Fig. 4B). We found that in the presence of potassium phthalimide, an OAC was formed that persisted for several hours [as determined by ^{19}F-nuclear magnetic resonance (^{19}F-NMR) characterization], whereas in absence of phthalimide, we obtained only protodehalogenation and reductive homocoupling products. The identity of the observed OAC was subsequently confirmed through independent synthesis.

Isolated complex ***3*** has substantial stability, which stands in stark contrast to the typically rapid decomposition of OACs lacking ortho substituents on the aryl ligand (*46*). Complex ***3*** is stable for at least 24 hours in the dark in both dimethyl sulfoxide (DMSO) and dichloromethane (DCM), as well as under irradiation for 2 hours in DMSO. Under reaction-relevant conditions, we found that decomposition was prevented when an excess of potassium phthalimide was added to prevent ligand exchange by carboxylates. Phthalimide thus likely precludes typical decomposition pathways by keeping complex ***3*** coordinatively saturated. Despite this inherent stability, complex ***3*** is fully able to capture radicals and undergo reductive elimination to form the desired cross-coupled product. Irradiation of complex ***3*** in the presence of an equimolar amount of $(Ir[dF(CF_3)ppy]_2(dtbpy))PF_6$ and excess of carboxylic acid, BTMG, and potassium phthalimide yields the desired cross-coupled product in 67% yield (see the supplementary materials).

We then set out to probe the importance of complex ***3*** to the reaction itself. Evaluation of the reaction progress of an activated acid (N-benzyloxycarbonyl-proline) and a nonactivated acid (cyclopentanoic acid) in the presence of phthalimide revealed that the nonactivated acid reacted more slowly overall (Fig. 4B, left). A Stern-Volmer analysis revealed a clear difference in the photocatalyst quenching rate between the two acids (Fig. 4B, middle). The slower quenching of nonactivated acids translates not only into slower formation of alkyl radicals but also slower formation of reduced iridium photocatalyst. By extension, we reasoned that this leads to slower Ni reduction and slower formation of the OAC. The lower steady-state concentration of alkyl radicals and OACs translates into a slower rate of radical capture by Ni, which allows for it to be outcompeted by the unimolecular OAC decomposition pathway. Stabilizing the OAC should therefore benefit nonactivated acids because it precludes the decomposition pathway and extends the time frame for successful radical capture. A PhotoNMR experiment (using cyclopentanoic acid) in the absence of phthalimide revealed no ^{19}F-NMR signals that could be assigned to an OAC. When this experiment was repeated in the presence of phthalimide, a new ^{19}F-NMR signal was observed that matched the signal obtained from complex ***3***, indicating a considerable increase in the steady-state concentration of the OAC (Fig. 4B, right). This further supports our hypothesis.

Along a different line of inquiry, we realized that electron-rich aryl bromides seemed to give particularly strong improvements in yield, and we suspected a correlation between the electronic properties of the aryl group and the effect of phthalimide. A Hammett study was conducted both in absence and in presence of phthalimide to probe this hypothesis (Fig. 4C, left). A moderately strong Hammett ρ of 1.57 was observed in the absence of phthalimide, which suggests that the oxidative addition of the aryl bromide contributes to the

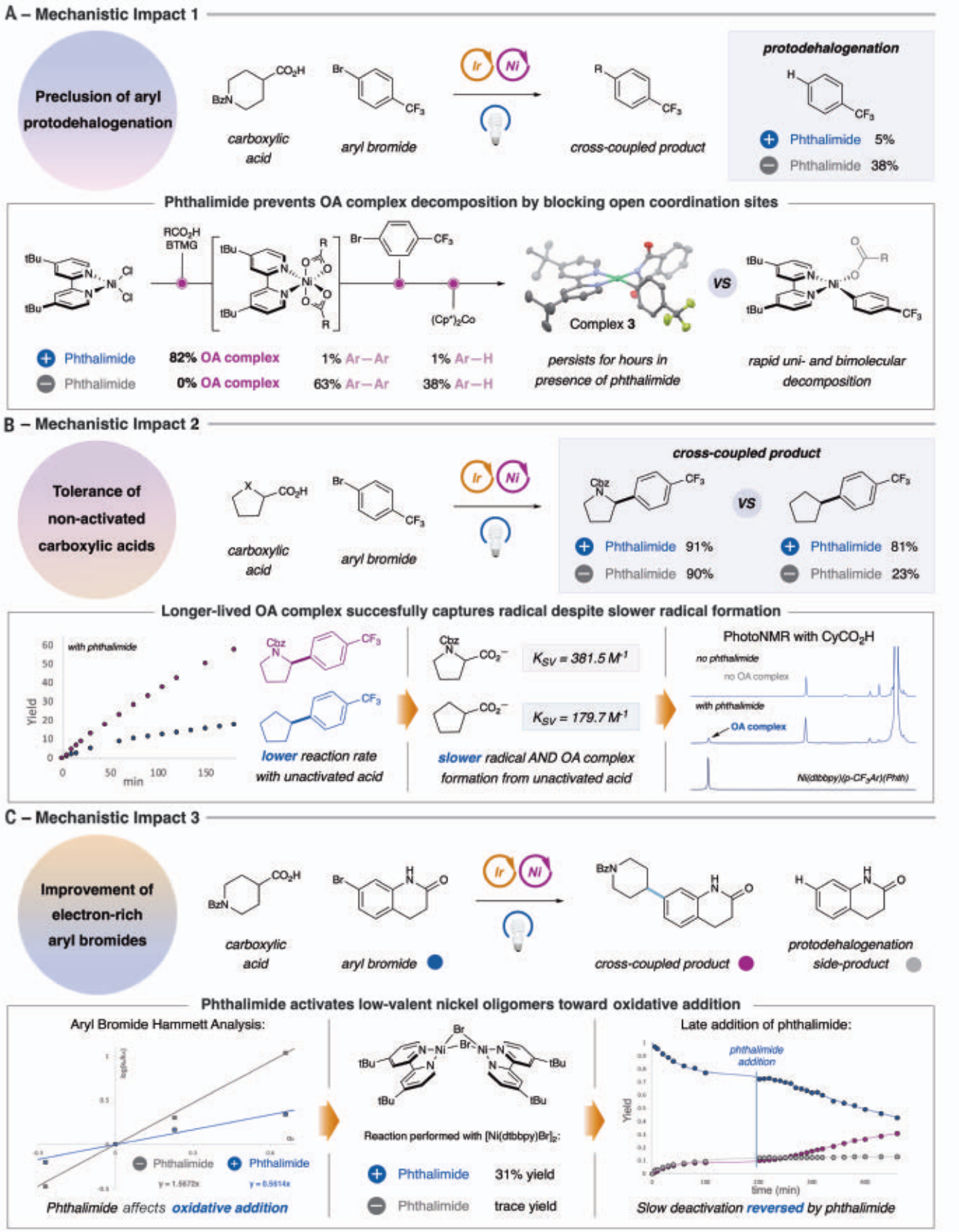


Fig. 4. Mechanistic insight. (**A**) Decrease of protodehalogenation observed with phthalimide caused by stabilization of the OAC. For the x-ray structure (50% probability ellipsoids), hydrogens and tBu/CF3 disorder are omitted for clarity. (**B**) Slower quenching is tolerated because of the increased OAC concentration. (**C**) Increasing active Ni on cycle by activation of low-valent Ni oligomers.

overall rate of the reaction. We found that the addition of phthalimide led to an overall increase in the initial reaction rate, in particular for electron-rich aryl bromides, and a lowered ρ of 0.56. We reasoned here that phthalimide must have some bearing on the oxidative addition.

To uncover the specific role of phthalimide in the oxidative addition, we evaluated the progress of a different model catalytic reaction in the presence and absence of phthalimide (Fig. 4C, top; substrate pair yields were 10% without phthalimide and 70% with phthalimide). Unexpectedly, in the absence of phthalimide, the reaction underwent deactivation over the first 100 min to reach an unproductive stationary state, whereas in the presence of phthalimide, the reaction steadily went to completion (figs. S24 and S25). In conjunction with the Hammett data, we found it reasonable to postulate a deactivation of the reaction caused by a progressive decrease in catalytically active Ni able to undergo oxidative addition. Phthalimide would thus act by bringing the Ni back on cycle.

To investigate the reactivation hypothesis, a competition study between phenyl bromide and various electronically distinct aryl bromides was conducted to assess product ratios at an early time point both in absence and presence of phthalimide (see the supplementary materials for details). For each aryl bromide, we found that product ratios did not change meaningfully when phthalimide was added (with the exception of the electron-poor 4-bromobenzotrifluoride). This is consistent with the hypothesis that phthalimide serves to increases the amount of Ni competent to undergo oxidative addition, which supports the proposed reactivation hypothesis.

It was reasonable to hypothesize that the deactivation could be attributed to the formation of low-valent Ni oligomers, which are known to be unreactive toward oxidative addition in their oligomeric form (*47*, *48*). Independently prepared dimer $[(dtbbpy)NiBr]_2$ was subjected to the reaction to test this hypothesis (Fig. 4C, middle). Trace yield of product was observed with the dimer for the previously used substrate pair, which confirms that this Ni species by itself is incompetent in this reaction. Repeating this experiment in the presence of phthalimide afforded the product in 31% yield, thus demonstrating that phthalimide can, at least in principle, return unreactive multimers into a monomeric, catalytically active state. Stable (bpy)Ni(I)-phthalimido complexes have been reported, which lends support to this hypothesis (*49*).

Taking the previous observations together, we reasoned that the addition of phthalimide to a deactivated reaction should lead to a reactivation of the catalysts and resumption of product formation. Phthalimide should break up any formed inactive multimeric species and thus allow for productive turnover. Consistent with our hypothesis, when phthalimide was doped into such a reaction, we observed almost complete reactivation and steady turnover (Fig. 4C, right).

Overall, we believe that phthalimide acts in two mechanistically distinct manners. First, phthalimide affects the stability of Ni-aryl complexes by acting as a ligand that precludes decomposition pathways such as protodehalogenation and aryl metathesis. This also accounts for the inclusion of nonactivated acids into the scope of this transformation, where the increased OAC lifetime counteracts the lower effective radical concentration inherent to these acids. Furthermore, when using electron-rich aryl bromides, the reaction undergoes reversible deactivation caused by a decrease in the effective concentration of on-cycle Ni catalyst. This is presumably because of the formation of off-cycle, unreactive multimeric species resulting from the aggregation of low-valent Ni complexes. Phthalimide is capable of reactivating inactive multimeric Ni species and thus increasing the concentration of catalytically active Ni, which allows the catalytic cycle to be productively turned over. We do not rule out further effects of phthalimide in this reaction, and additional mechanistic investigations are currently ongoing. The effect of phthalimide on Ni-catalyzed cross-couplings in general is also undergoing further investigation.

In summary, by identifying phthalimide as a beneficial additive for decarboxylative arylation, we were able to develop a general transformation to reliably form $C(sp^2)$–$C(sp^3)$ bonds from feedstock chemicals. This improvement was achieved in less than a year from project inception, which highlights the expedited nature of additive mapping. Furthermore, we then leveraged the discovered phthalimide effect to broaden the mechanistic understanding of Ni-catalyzed cross-coupling from an angle not accessible by traditional mechanistic studies, which highlights the orthogonality of this approach. We imagine that both the use of phthalimide as an additive and the herein reported approach for reaction generalization and mechanistic elucidation will be rapidly embraced to affect modern organic synthesis.

REFERENCES AND NOTES

1. T. M. Trnka, R. H. Grubbs, *Acc. Chem. Res.* **34**, 18–29 (2001).
2. N. Miyaura, A. Suzuki, *Chem. Rev.* **95**, 2457–2483 (1995).
3. P. Ruiz-Castillo, S. L. Buchwald, *Chem. Rev.* **116**, 12564–12649 (2016).
4. H. C. Brown, *Hydroboration* (W. A. Benjamin, 1962)
5. J. Tsuji, *Palladium Reagents and Catalysts* (Wiley, 1995)
6. C. C. C. Johansson Seechurn, M. O. Kitching, T. J. Colacot, V. Snieckus, *Angew. Chem. Int. Ed.* **51**, 5062–5085 (2012).
7. Y. Park, Y. Kim, S. Chang, *Chem. Rev.* **117**, 9247–9301 (2017).
8. W. L. Truett, D. R. Johnson, I. M. Robinson, B. A. Montague, *J. Am. Chem. Soc.* **82**, 2337–2340 (1960).
9. T. L. Choi, R. H. Grubbs, *Angew. Chem. Int. Ed.* **42**, 1743–1746 (2003).
10. A. S. Guram, S. L. Buchwald, *J. Am. Chem. Soc.* **116**, 7901–7902 (1994).
11. F. Paul, J. Patt, J. F. Hartwig, *J. Am. Chem. Soc.* **116**, 5969–5970 (1994).
12. B. T. Ingoglia, S. L. Buchwald, *Org. Lett.* **19**, 2853–2856 (2017).
13. R. T. Conlin, P. P. Gaspar, R. H. Levin, M. Jones Jr., *J. Am. Chem. Soc.* **94**, 7165–7166 (1972).
14. K. Liao, S. Negretti, D. G. Musaev, J. Bacsa, H. M. L. Davies, *Nature* **533**, 230–234 (2016).
15. J. A. Lee, M. T. Uhlik, C. M. Moxham, D. Tomandl, D. J. Sall, *J. Med. Chem.* **55**, 4527–4538 (2012).
16. J. G. Moffat, F. Vincent, J. A. Lee, J. Eder, M. Prunotto, *Nat. Rev. Drug Discov.* **16**, 531–543 (2017).
17. B. K. Wagner, S. L. Schreiber, *Cell Chem. Biol.* **23**, 3–9 (2016).
18. D. G. Brown, H. J. Wobst, *J. Med. Chem.* **63**, 1823–1840 (2020).
19. W. Zheng, N. Thorne, J. C. McKew, *Drug Discov. Today* **18**, 1067–1073 (2013).
20. D. C. Swinney, J. Anthony, *Nat. Rev. Drug Discov.* **10**, 507–519 (2011).
21. Q. T. L. Pasquer, I. A. Tsakoumagkos, S. Hoogendoorn, *Molecules* **25**, 5702–5724 (2020).
22. W. Dejonghe, E. Russinova, *Plant Physiol.* **174**, 5–20 (2017).
23. M. Szabo *et al.*, *Drug Des. Devel. Ther.* **11**, 1957–1967 (2017).
24. A. L. Casado, P. Espinet, A. M. Gallego, *J. Am. Chem. Soc.* **122**, 11771–11782 (2000).
25. J. C. Vantourout, H. N. Miras, A. Isidro-Llobet, S. Sproules, A. J. B. Watson, *J. Am. Chem. Soc.* **139**, 4769–4779 (2017).
26. Z. Dong, J. Wang, G. Dong, *J. Am. Chem. Soc.* **137**, 5887–5890 (2015).
27. A. Bellomo *et al.*, *Angew. Chem. Int. Ed.* **51**, 6912–6915 (2012).
28. P. J. Sarver *et al.*, *Nat. Chem.* **12**, 459–467 (2020).
29. F. Lovering, J. Bikker, C. Humblet, *J. Med. Chem.* **52**, 6752–6756 (2009).
30. W. P. Walters, J. Green, J. R. Weiss, M. A. Murcko, *J. Med. Chem.* **54**, 6405–6416 (2011).
31. R. Jana, T. P. Pathak, M. S. Sigman, *Chem. Rev.* **111**, 1417–1492 (2011).
32. C. Han, S. L. Buchwald, *J. Am. Chem. Soc.* **131**, 7532–7533 (2009).
33. N. Fleury-Brégeot *et al.*, *J. Org. Chem.* **77**, 10399–10408 (2012).
34. L. Li, S. Zhao, A. Joshi-Pangu, M. Diane, M. R. Biscoe, *J. Am. Chem. Soc.* **136**, 14027–14030 (2014).
35. J. H. Kirchhoff, M. R. Netherton, I. D. Hills, G. C. Fu, *J. Am. Chem. Soc.* **124**, 13662–13663 (2002).
36. I. D. Hills, M. R. Netherton, G. C. Fu, *Angew. Chem. Int. Ed.* **42**, 5749–5752 (2003).
37. A. W. Dombrowski *et al.*, *ACS Med. Chem. Lett.* **11**, 597–604 (2020).
38. Z. Zuo *et al.*, *Science* **345**, 437–440 (2014).
39. J. Hioe, H. Zipse, *Org. Biomol. Chem.* **8**, 3609–3617 (2010).
40. R. W. Evans, *Novel Coupling Reactions Utilizing Base Metal and Photoredox Catalysis* (Princeton Univ. Press, 2017).
41. K. D. Collins, F. Glorius, *Nat. Chem.* **5**, 597–601 (2013).
42. S. D. Dreher, S. W. Krska, *Acc. Chem. Res.* **54**, 1586–1596 (2021).
43. N. Vervoort, D. Daemen, G. Török, *J. Chromatogr. A* **1189**, 92–100 (2008).
44. Products were identified using mass spectrometry and UV210 and UV250 signatures. Selected examples were scaled up to confirm product formation (see the supplementary materials).
45. K. D. Dykstra *et al.*, *ACS Med. Chem. Lett.* **12**, 337–342 (2021).
46. T. Yamamoto, S. Wakabayashi, K. Osakada, *J. Organomet. Chem.* **428**, 223–237 (1992).
47. N. A. Till, S. Oh, D. W. C. MacMillan, M. J. Bird, *J. Am. Chem. Soc.* **143**, 9332–9337 (2021).
48. R. Sun *et al.*, *J. Am. Chem. Soc.* **141**, 89–93 (2019).
49. B. Nestler, E. Uhlig, *Z. Anorg. Allg. Chem.* **530**, 196–206 (1985).

ACKNOWLEDGMENTS

We thank N. Till, P. Sarver, and Z. Dong for helpful discussions; T. Liu for support with graphic design; and P. Jeffrey for assistance with crystallography. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences of the National Institutes of Health (NIH). **Funding:** This work was supported by the National Institute of General Medical Sciences of the National Institutes of Health (award R35GM134897-02) and by Merck Sharp & Dohme Corp., a subsidiary of Merck & Co., Inc.; Kenilworth, NJ, USA; Pfizer; Janssen; Bristol-Myers Squibb; Genentech; and Celgene. C.N.P.K. was supported by Princeton University, E. Taylor, and the Taylor family by an Edward C. Taylor fellowship. J.A.K. was supported by a predoctoral fellowship (award DGE 1656466) from the National Science Foundation. **Author contributions:** S.W.K., S.D.D., and D.W.C.M. conceived of the work. All authors designed the experiments. C.N.P.K., J.A.K., T.N., and S.D.D. performed and analyzed the experiments. C.N.P.K., J.A.K., S.W.K., S.D.D., and D.W.C.M. prepared the manuscript. **Competing interests:** D.W.C.M. declares a competing financial interest with respect to the integrated photoreactor (commercialized as PennPhD Photoreactor M2). The remaining authors declare no competing interests. **Data and materials availability:** Data are available in the supplementary materials. The Cambridge Crystallographic Data Centre (CCDC) website contains the supplementary crystallographic data for this paper, which can be obtained free of charge at www.ccdc.cam.ac.uk/data_request/cif under identifier 2153752.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abn1885
Materials and Methods
Supplementary Text
Figs. S1 to S35
NMR Spectral Data
References (*50*–*55*)
Excel Data for HTE Experiments
Data for Experiments 9.1 and 10.1 to 10.3

9 November 2021; accepted 25 March 2022
10.1126/science.abn1885

GLOUCESTER MARINE GENOMICS INSTITUTE

CHIEF SCIENTIFIC OFFICER

Gloucester Marine Genomics Institute seeks candidates for the position of Chief Scientific Officer based in Gloucester, MA

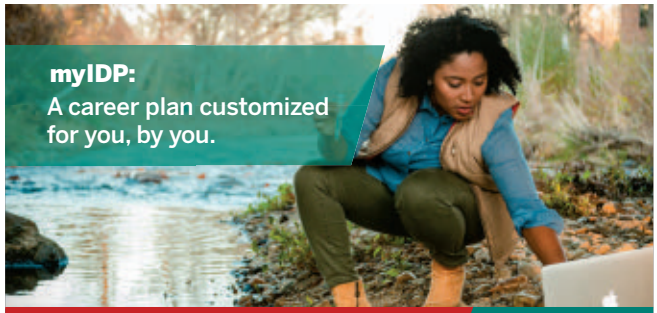
GMGI has increased the prominence of its scientific research in marine genomics. Having established a world-class research program over the last eight years, GMGI is positioned for substantial growth and seeks a visionary scientific leader to lead the organization's research with an integrated strategic vision for long-term success, and enable the recruitment of exceptional junior and senior research scientists. The CSO will consider the interplay of GMGI's expertise in marine genetics and genomics, its research, educational and commercial strengths, as well as the evolving ecosystem of biomedical research. While honoring Gloucester's unique culture and history, the CSO will drive the organization forward with insight and creativity to a bold vision of its future, while also leading with clarity and a strong "human" touch.

The CSO will be responsible for the Institute's overall scientific direction, establishing key strategic goals and setting organizational performance expectations. The successful candidate will have an international reputation and well-funded research program. They will create the necessary environment, processes and infrastructure to discuss and make critical strategic, financial and scientific decisions.

GMGI addresses critical challenges facing our oceans, human health and the environment through innovative scientific research and education. By bringing world-class science and transformative workforce development to Gloucester's historic waterfront, GMGI is catalyzing the regional economy. Founded in 2013, our work includes a state-of-the-art research institute where we conduct marine research powered by genomics, and the Gloucester Biotechnology Academy, preparing young adults to become biotech lab technicians.

GMGI is proud to be an equal opportunity employer, and all aspects of employment are based on merit, competence, performance, and business need.

Interested individuals submit materials to: recruiting@gmgi.org



For your career in science, there's only one Science

Features in myIDP include:

- Exercises to help you examine your skills, interests, and values.
- A list of 20 scientific career paths with a prediction of which ones best fit your skills and interests.

myIDP Visit the website and start planning today!
myIDP.sciencecareers.org

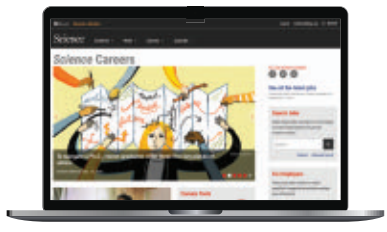
ScienceCareers In partnership with:

FASEB · University of Massachusetts Medical School UMASS · UCSF University of California San Francisco · UIC · BURROUGHS WELLCOME FUND

Science Careers helps you advance your career. Learn how!

- Register for a free online account on **ScienceCareers.org**.
- Search hundreds of job postings and find your perfect job.
- Sign up to receive e-mail alerts about job postings that match your criteria.
- Upload your resume into our database and connect with employers.
- Watch one of our many webinars on different career topics such as job searching, networking, and more.
- Download our career booklets, including Career Basics, Careers Beyond the Bench, and Developing Your Skills.
- Complete an interactive, personalized career plan at "my IDP."
- Visit our Employer Profiles to learn more about prospective employers.
- Read relevant career advice articles from our library of thousands.

Visit **ScienceCareers.org** today — all resources are free



SCIENCECAREERS.ORG

ScienceCareers
FROM THE JOURNAL SCIENCE AAAS

北京清华长庚医院
Beijing Tsinghua Changgung Hospital

Tel: +86-10-56118687 **Email:** medtalent@mail.tsinghua.edu.cn

Beijing Tsinghua Changgung Hospital (BTCH), established on November 28, 2014 as a Tsinghua University-affiliated public hospital, is a world-class academic medical center which integrates services, training and research alongside the mission to promote health, cultivate outstanding physicians, and lead medical innovations while maintaining excellent management. BTCH hosts 7 medical departments which include 52 specialized disciplines, offering comprehensive services from internal medicine to surgery, obstetrics-gynecology, pediatrics, emergency and intensive Care and family medicine. In addition to these, the hospital has built renowned integrated centers including Hepato-Pancreato-Biliary Center, Neuro-Center, Digestive Disease Center, Emergency and Intensive Care Center and Cardiovascular Center. Our medical professionals are from prestigious backgrounds and provide diverse perspectives to medical practices internationally. This team is led by the inaugural president Jiahong Dong, an academician of the Chinese Academy of Engineering and an internationally recognized HPB and transplant surgeon.

BTCH aims to provide health-oriented medical services which are guided by key tenets which give first priority to patients and operate with diligent doctors as its core foundation. We conceptualized the idea of precision healthcare, compassionate service and lean management, in order to visualize a modern healthcare service system which merges medical prevention, treatment, rehabilitation, chronic illness management and senior healthcare. Through an accelerated 7-year development, BTCH has become a comprehensive high-quality hospital with good hospital governance, a strong professional team, a pool of outstanding talents, first-class medical quality and a good reputation in the industry.

Our services are grounded in cutting-edge medical technologies, evidence-based practices and standardized care protocols with value-based care. Abiding by the belief of " Perform surgeries that others are unwilling to risk, and cure diseases that others are incapable of", our hospital has successfully treated a great number of complicated and difficult cases from home and abroad by confronting medical problems in various specialized fields. Furthermore, BTCH is dedicated to establishing a regional integrated healthcare consortium focusing on the health of community residents, by consistently providing hierarchical, integrated medical care which encompasses the whole population, all aspects of healthcare and the one's entire lifespan.

As an initiative undertaken by BTCH, the Institute of Precision Medicine of Tsinghua University serves as an interdisciplinary health science innovation platform, supporting 6 medical-engineering research centers and 15 interdisciplinary medical laboratories including those of big data, digital health, intelligent healthcare, medical robotics, wearable healthcare device, ECMO, translational medical imaging, regional integrated medical systems, photomedicine, liver metabolomics, precision pharmacy, etc.

Guided by the precision healthcare theory, the Clinical-Driven Research (CDR) paradigm and the Bedside to Bench to Clients (BBC) model was established to combine Tsinghua University's academic resources in clinical medicine, science, engineering, public health, arts and others for multidisciplinary collaboration. This platform accumulates into an innovation ecosystem that bridges healthcare organizations, the academia and enterprises, through translating basic and clinical research into better and safer healthcare practices, and implementation of products, a number of which are currently either in the later stages of clinical trials or in use.

BTCH devotes to cultivating outstanding future physician leaders through the incorporation of benevolent healthcare practices, scientific aptitude and leadership. The hospital's accredited residency and fellowship programs are certified by the National Health Commission of China with yearly assessments from rigorous national level examinations yielding scores ranking in the top 5th percentile. The hospital was awarded first place in the 2019 annual national residency training examination as well as designated as one of the national clinical training demonstration centers in 2018. Our system-oriented training programs foster health professionals that adapt well to the changing landscape of healthcare.

In following our "4I" development strategy that highlights Innovation, Integration, individuation and Internationalization, we embrace talented physicians, surgeons, scientists and other medical professionals of all nationalities and backgrounds to join our BTCH family. We are committed to providing a superior platform, competitive benefits and comfortable working environment for each faculty member. Let's convene with the common mission of creating a world class center of excellence, cultivating outstanding physicians, provoking cutting-edge medical innovations and achieving " Healthy China", a healthful shared future for mankind.

ADVERTISEMENT

By **Asma Bashir**

Scientist and artist

"Would you say you are more of a scientist or an artist?" I frequently heard this question as I was preparing to start my neuroscience doctoral program, often after I told scientific mentors and peers that I love writing poetry. I learned to respond, "A scientist"—not because it was necessarily true, but because I sensed it would win me more respect. As an undergraduate, poetry had been my therapy, my cathartic release. But in grad school, I wanted to be seen as just as dedicated to my work as those around me, who appeared completely undistracted by nonscientific pursuits. So, I left my poetry by the wayside—but it wasn't long before I felt a massive void in my life.

By the second year of my Ph.D., my thesis project had become all-consuming and I was starting to burn out. One afternoon, an undergraduate trainee who worked in the lab across the hall mentioned the weekly event held by the university's poetry club, of which she was a leader. After hearing how much I enjoyed poetry, she declared, "You should come!" A few days later, I sat in a packed café, exulting in the warmth of the vibrant community.

When I received another invitation—this time to perform my own work—I couldn't say no. I rediscovered catharsis through storytelling, writing about workplace misogyny I had witnessed. As I prepared my poem, though, I was also awash with guilt about dedicating hours to writing and rehearsing, time I could have spent reading papers and analyzing data. Still, I finished my poem and showed up for my performance. With the outpouring of love I received that evening, I knew I'd be back.

I continued to write and perform, but I still worried about appearing less than fully dedicated to my science. So, when a senior grad student in my lab whispered, "I hear you perform poetry!" I froze. Was my cover blown? My heart began to race. She went on to invite me to perform at our department's art gala—a talent show of sorts, showcasing faculty, staff, and trainees. "By the way, Cheryl will be performing!" she added.

I was surprised to discover that Cheryl, our lab's principal investigator, was a dancer. I excitedly sought her out, and we spoke for more than an hour about how science and art can intertwine. She explained that my skill at picking up patterns in experimental results was the same one I used to create rhymes in my poems. With Cheryl as a model, I decided to perform at the gala, for once not feeling I had anything to hide.

"I decided to perform at the gala, for once not feeling I had anything to hide."

The night of the gala, I marveled at the talent on display, including stunning musicianship and visual art. Chatting with the other performers, many of whom were also graduate students, I realized many shared my concern that peers might interpret a passion for art as a lack of passion for science or, at the very least, a distraction. But over the course of the evening, I found myself at peace with the two halves of my identity, and I got the feeling others felt that same peace, too.

Since then, I have continued to embrace all of me. While completing my Ph.D. and moving on to a postdoc, I frequently scribbled lines of poetry on Post-its and Kimwipes and wrote poems during my lunch break. Some mentors and peers have seen the value of this outside interest; others have been skeptical. But when I have my own doubts, I think of my Ph.D. supervisor, and of renowned 19th century neuroscientist Santiago Ramón y Cajal, whose beautiful drawings of the central nervous system can be found in science textbooks and art museums in equal measure. They remind me that art and science can be profoundly complementary. And now, if someone asks me whether I'm more of a scientist or an artist, I very proudly say, "Both." ■

Asma Bashir is a neuroscientist by training and host of the podcast Her Royal Science in Halifax, Canada. Send your career story to SciCareerEditor@aaas.org.

ILLUSTRATION: ROBERT NEUBECKER

CALL FOR PAPERS

SPJ
SCIENCE PARTNER JOURNALS

PEKING UNIVERSITY 1898

Health Data Science

Health Data Science is an online-only Open Access Science Partner Journal published in affiliation with **Peking University (PKU)** and distributed by the **American Association for the Advancement of Science (AAAS)**. This journal is committed to promoting the appropriate application of cutting-edge technologies and analytic approaches for the health field, translating the value of health data into practice for the benefit of patients' health and populations' well-being and advancing health data science. *Health Data Science* publishes innovative, scientifically-rigorous original research, comprehensive reviews, insightful perspectives, viewpoints and editorials. The topics of interest include but are not limited to: data management, governance and provenance; comparative effectiveness research; ethics issues; public health and health policies; and more.

Submit your research to *Health Data Science* today!

Learn more at **spj.sciencemag.org/hds**

The Science Partner Journal (SPJ) program was established by the American Association for the Advancement of Science (AAAS), the nonprofit publisher of the *Science* family of journals. The SPJ program features high-quality, online-only, Open Access publications produced in collaboration with international research institutions, foundations, funders and societies. Through these collaborations, AAAS furthers its mission to communicate science broadly and for the benefit of all people by providing top-tier international research organizations with the technology, visibility, and publishing expertise that AAAS is uniquely positioned to offer as the world's largest general science membership society.

Visit us at **spj.sciencemag.org**

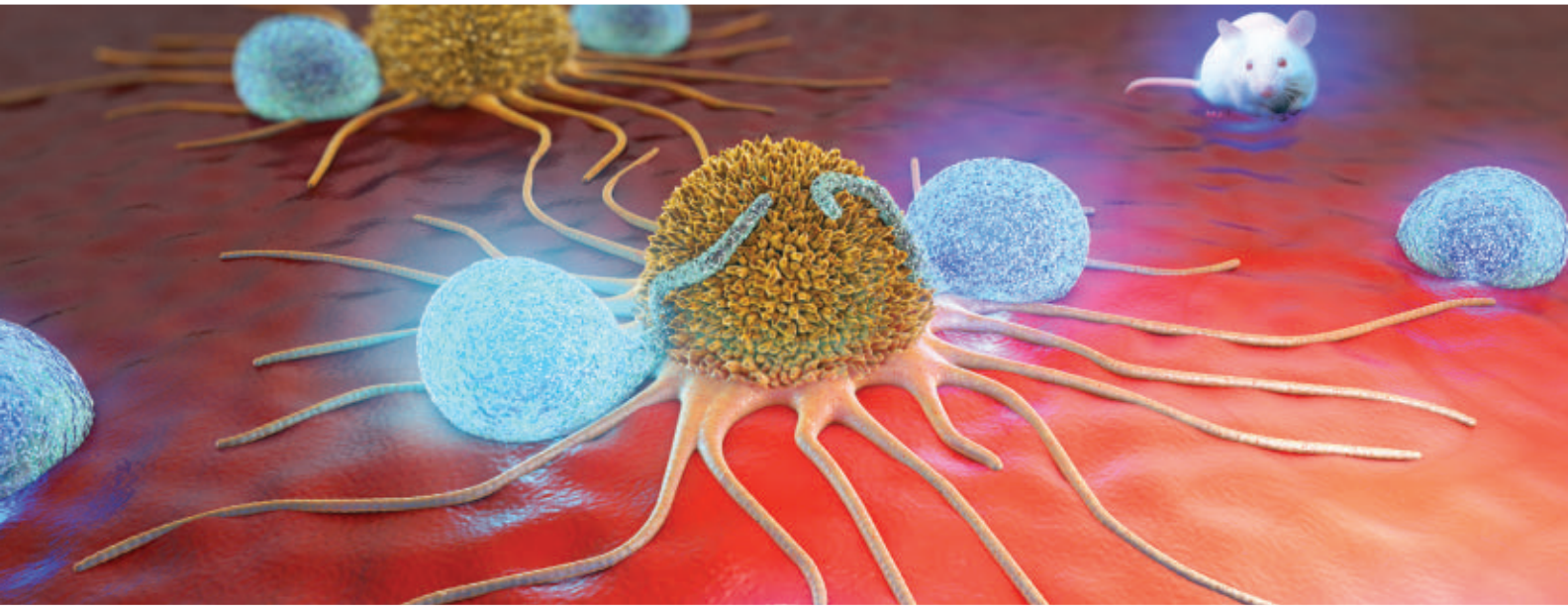
@SPJournals @SPJournals

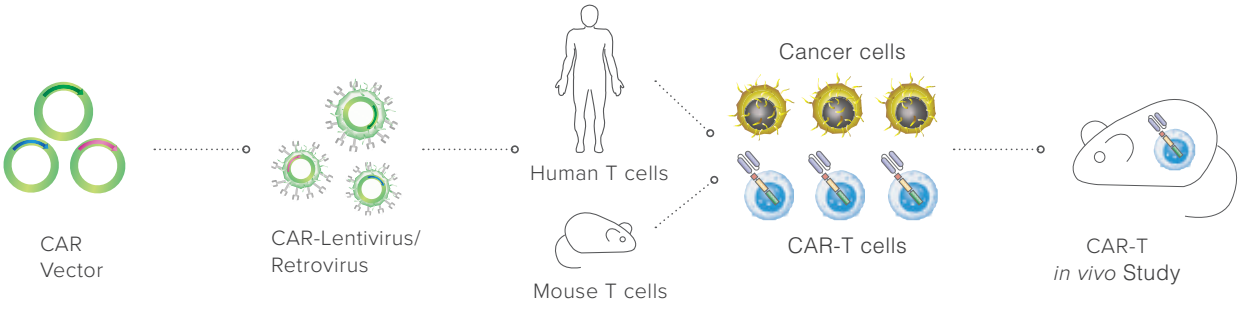
OPEN ACCESS

ARTICLE PROCESSING CHARGES WAIVED THROUGH DECEMBER 2024

CAR-T Animal Studies

Testing *In Vivo* Efficacy & More for Cell Therapy Research





Animal Research Custom Services

Custom services available.

- IVIS imaging
- Toxicology
- PK/PD models
- Immunohistochemistry (IHC)
- Cytokine panels
- Tumor volume and weight
- Histopathology
- Genomic parameters
- Cell phenotyping

ProMab Biotechnologies offers a fully comprehensive in vivo functionality service to analyze your cell therapy product using xenograft NSG mice models. CAR-T cells can be made to expedite preclinical research studies in xenogeneic or syngeneic mouse models. ProMab has developed patient-derived xenograft models of different types of cancers to study and validate the efficacy of cell therapy products against solid and hematological tumors and its microenvironment.

All products are for research only

Discover more | www.promab.com

ProMab Biotechnologies, Inc.
2600 Hilltop Dr, Building B, Suite C320, Richmond, CA 94806
1.866.339.0871 | info@promab.com

SCAN ME